Article

Use of Hypoxic Respiratory Challenge for Differentiating Alzheimer's Disease and Wild-Type Mice Non-Invasively: A Diffuse Optical Spectroscopy Study

Myeongsu Seong [1,2] **, Yoonho Oh** [3] **, Hyung Joon Park** [4,5] **, Won-Seok Choi** [4,*] **and Jae Gwan Kim** [3,*]

1 School of Information Science and Technology, Nantong University, Nantong 226019, China
2 Research Center for Intelligent Information Technology, Nantong University, Nantong 226019, China
3 Department of Biomedical Science and Engineering, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea
4 School of Biological Sciences and Technology, College of Natural Sciences, College of Medicine, Chonnam National University, Gwangju 61186, Republic of Korea
5 Department of Biochemistry, University of Washington, Seattle, WA 98195, USA
* Correspondence: choiw@chonnam.ac.kr (W.-S.C.); jaekim@gist.ac.kr (J.G.K.)

Abstract: Alzheimer's disease is one of the most critical brain diseases. The prevalence of the disease keeps rising due to increasing life spans. This study aims to examine the use of hemodynamic signals during hypoxic respiratory challenge for the differentiation of Alzheimer's disease (AD) and wild-type (WT) mice. Diffuse optical spectroscopy, an optical system that can non-invasively monitor transient changes in deoxygenated (ΔRHb) and oxygenated (ΔOHb) hemoglobin concentrations, was used to monitor hemodynamic reactivity during hypoxic respiratory challenges in an animal model. From the acquired signals, 13 hemodynamic features were extracted from each of ΔRHb and $-\Delta OHb$ (26 features total) for more in-depth analyses of the differences between AD and WT. The hemodynamic features were statistically analyzed and tested to explore the possibility of using machine learning (ML) to differentiate AD and WT. Among the twenty-six features, two features of ΔRHb and one feature of $-\Delta OHb$ showed statistically significant differences between AD and WT. Among ML techniques, a naive Bayes algorithm achieved the best *accuracy* of 84.3% when whole hemodynamic features were used for differentiation. While further works are required to improve the approach, the suggested approach has the potential to be an alternative method for the differentiation of AD and WT.

Keywords: Alzheimer's disease; diffuse optical spectroscopy; hemodynamic measurements; hypoxic respiratory challenge; machine learning

1. Introduction

Dementia is one of the most devastating brain diseases and has various causes. The most common cause of dementia is Alzheimer's disease (AD) [1]. The prevalence of AD is expected to keep increasing as life expectancy increases. Generally, AD is characterized by the accumulation of β-amyloid and tau protein in the brain [1–5]. While AD results in severe damage to the brain, which deteriorates the life of the patient and the patient's family, there is no viable cure for AD at present. However, early diagnosis of AD is still beneficial for future planning, participation in clinical trials, and delaying the progress of the disease by regulating the patient's lifestyle [1]. For the diagnosis of AD, various methods are used. The most conventional diagnostic methods are cognitive and behavioral tests [1]. Other than cognitive and behavioral tests, blood sampling and brain imaging have been used to diagnose AD. Among imaging methods, monitoring of β-amyloid [2,3] and tau protein [4,5] using positron emission tomography (PET) has been performed. Blood oxygenation level-dependent (BOLD)-functional magnetic resonance imaging (fMRI), a

type of MRI-based technique, has also been used to investigate the differences between control, mild cognitive impairment (MCI), and AD [6]. However, the use of PET and MRI is usually costly. As an alternative approach, optical techniques, solely or combined with other techniques, have been utilized to characterize AD. Among various optical techniques, diffuse optical spectroscopy (DOS), also known as near-infrared spectroscopy (NIRS) in biomedical engineering, is frequently used in AD studies due to its relatively high temporal resolution, low cost, portability, and non-invasiveness. In most NIRS-based AD studies, experimental protocols, including the verbal fluency test [7–12], N-back test (including 1-back) [10–13], oddball test [11,12], and Stroop test [10], are used as stimuli to the brain. Such approaches aim to test neurovascular coupling to hypothesize differences in the responses of an AD group and a normal group. However, the applicability of such experimental protocols may be affected by the educational level of an individual.

Monitoring the cerebrovascular reactivity (CVR) caused by supplying breathing gas or using specific breathing methods is another approach to investigating various brain diseases, including migraine [14], traumatic brain injury [15], epilepsy [16], and cerebral small vessel disease [17]. In the meantime, some animal studies have used breathing gas challenges supplying 100% O_2 gas or carbogen gas (95% O_2 gas balanced by 5% CO_2 gas) to differentiate between 3xTg, a type of AD animal model, and wild-type (WT) animals [18,19]. However, these studies did not test the use of breath-holding as a means to induce the CVR. Similarly, a spatial frequency domain imaging system, which is used in these studies, requires excision of the scalp, which cannot be easily performed on human subjects.

While many AD studies using optical techniques performed relatively simple analyses including statistical tests and comparison of signal correlations [7–9,11,13], thanks to the rapid development of artificial intelligence (AI) techniques such as machine learning (ML) and deep learning (DL), recent AD studies using optical techniques tend to perform more advanced analyses of acquired signals [10,12]. It is well known that AI techniques extract 'hidden' information that cannot be found by conventional analytical methods, thus it is believed that a wide range of medical fields will benefit from adopting AI techniques.

In this work, to explore an alternative non-invasive, simple, and cost-effective approach for diagnosing AD, preliminary animal experiments were performed during the hypoxic respiratory challenge, an intervention that mimics breath-holding. During the experiments, a DOS system was used to monitor the CVR of AD and WT. In the meantime, Monte Carlo simulations were performed to show the possibility of non-invasive brain signal measurement using the DOS system. Statistical analysis and machine learning (ML) algorithms were used to confirm the feasibility of the suggested approach and the practicality of the protocol for differentiating AD and WT.

Contributions of this study include: (1) testing hypoxic respiratory challenge in AD and WT by measuring hemodynamic signals using the DOS system; (2) showing the usefulness of extracted hemodynamic features, which are based on methods suggested in previous breast cancer studies, in the differentiation of AD and WT; (3) identifying some hemodynamic features that can differentiate AD and WT statistically significantly; (4) testing 15 ML algorithms for hemodynamic features of CVR caused by hypoxic respiratory challenge for classification of AD and WT.

2. Materials and Methods

2.1. Animal Model and Preparation

We used 12-month-old male 5xFAD animals (AD, $n = 6$) and age-matched wild-type animals (WT, $n = 6$) in this study. The 5xFAD model is a genetically-modified mouse model that co-expresses 5 mutations of familial Alzheimer's disease (FAD) [20–23], and develops accumulation of β-amyloid in a relatively short time. As a result, the model shows the characteristics (amyloid pathology and memory impairment) of FAD in a shorter time in comparison to other Alzheimer's models. To induce anesthesia, each animal was initially placed in an induction chamber filled with 5% isoflurane mixed with 50% O_2 gas balanced by N_2 gas. After induction, the anesthesia was maintained with 1.5% isoflurane

supplied through a nose cone (Figure 1a). The mouse head was fixed using a stereotaxic frame (Customized stereotaxic frame, Digitaxis). To prevent eye dryness, an eye ointment was applied to both eyes of the animal. The fur on the head and the thigh was shaved and depilated to minimize the influence of scattering in the DOS and pulse oximeter measurements, respectively.
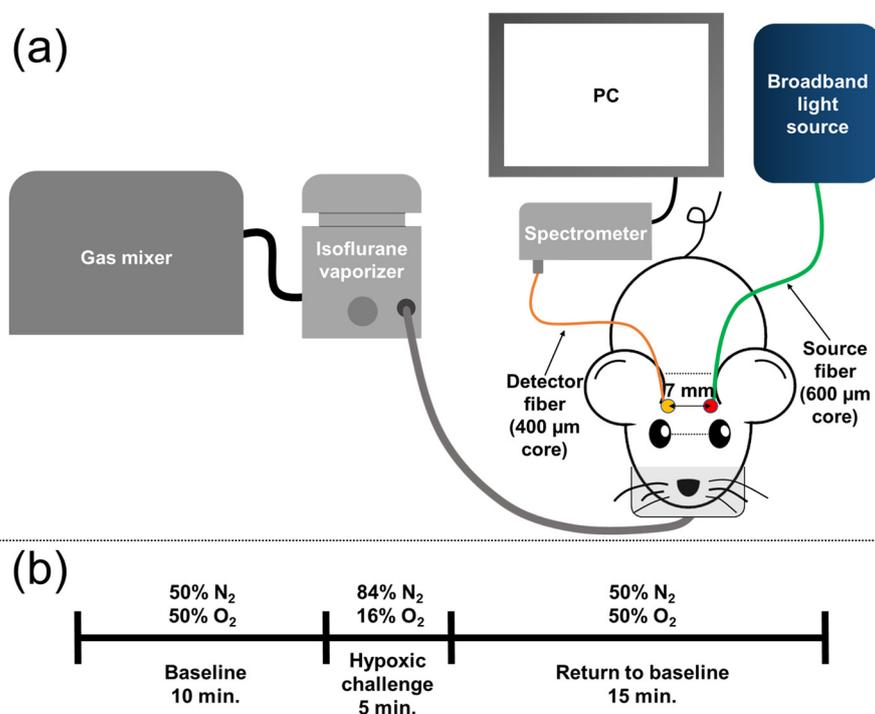
Figure 1. (**a**) Schematic of the animal experimental setup, (**b**) protocol of gas intervention.

2.2. Behavioral Tests

Before proceeding with the measurements of cerebral hemoglobin concentration, novel object recognition and Morris water maze tests were performed to confirm the AD-related behavioral changes in 5xFAD mice when the animals were 6–7 months old (5xFAD: $n = 6$; WT: $n = 6$).

2.2.1. Novel Object Recognition Test

The novel object recognition test consists of training and test phases. Mice were placed for 10 min in the open-top box ($40 \times 40 \times 40$ cm^3) where two identical objects were located during the training phase. The test phase started 24 h after the training phase. Mice were returned to the box with one novel object and the identical object used in the training phase and then explored the objects for 10 min. The preference index was calculated as the time to explore the novel object, which is expressed as 'a', divided by the time to explore both objects, which is expressed as 'a + b' (% preference index = {a/(a + b)} $\times$ 100 (%)).

2.2.2. Morris Water Maze Test

The Morris water maze test was performed over 4 trials per day for 4 consecutive days to test spatial memory. A circular test pool (114 cm in diameter, 25 cm in height) was filled with opaque water using non-toxic paint and was maintained at 25 °C. For all sessions over 4 days, a clear escape platform (15×10.5 cm^2) was immersed 1.5 cm below the water surface in a fixed position in the pool. In each trial, mice were randomly placed facing the wall in one of the three quadrants of the pool except for the quadrant containing the escape platform. Mice were allowed to swim to find the escape platform for 60 s. If the mice could not find the escape platform within 60 s, we guided the mice to the platform. The animals

were allowed to stay on the platform for 10 s. The latency time from the starting point to the escape platform was measured for all trials.

2.3. Diffuse Optical Spectroscopy for Cerebral Hemoglobin Concentration Measurement

Figure 1 shows a schematic of the system. The DOS system consisted of a broadband light source (HL-2000-HP, Ocean Optics, Orlando, FL, USA) with a 600-μm multi-mode fiber (M29L02, numerical aperture (N.A.): 0.39, Thorlabs, Newton, NJ, USA) and a spectrometer (custom USB4000, wavelength: 470–1000 nm, Ocean Optics, Orlando, FL, USA) with a 400-μm multi-mode fiber (M79L01, N.A.: 0.39, Thorlabs, Newton, NJ, USA). A source-detector separation was set to 0.7 cm to balance the signal stability, which was determined by comparing the maximum intensity and the noise level, as well as the signal acquisition rate, which is inversely proportional to the integration time. The expected probing depth of the system was approximately 0.35 cm based on the diffusion theory, which was deep enough to reach the depth of the brain. Even though the overall signal-to-noise ratio may increase as the source-detector separation becomes shorter, a shorter separation is not considered due to the multi-layer nature of the brain. Specifically, the decrease of the source-detector separation will enhance the signal sensitivity on the extracerebral tissues, including the scalp and the skull, and simultaneously diminish the signal sensitivity on the brain. Thus, the source-detector separation should be long enough to achieve high enough sensitivity on the CVR. The validity of the acquisition of the signals from the brain is explored using Monte Carlo simulation (Sections 2.6 and 3.2). DOS data were collected using a bundle software from Ocean Optics (Spectrasuite, Ocean Optics, Orlando, FL, USA). The sampling frequency of DOS varied depending on the integration time of the spectrometer since the animals showed different peak signal intensities. The integration time was set to have the maximum intensity value 3 or 4 times higher than the dark signal intensity, which resulted in 1 Hz to 3 Hz of the sampling frequency. The collected DOS raw data were saved for offline processing.

2.4. Signal Acquisition

The measurement was performed non-invasively, thus the bregma-lambda coordinate points could not be used. Instead, the DOS probe was placed between the horizontal line of the eyes and the horizontal line of the ears to acquire DOS signals from the same position. The probe was placed on the head of the mouse by the same person so that the position and pressure given to the head by the probe could be as constant as possible. Figure 1a shows the orientation of the source and detector fibers. Figure 1b shows the experimental protocol of the breathing gas challenge used in this work. The total measurement took 30 min. We supplied 50% O_2 gas balanced by N_2 gas for 10 min for the baseline measurement. After the baseline measurement, 16% O_2 gas balanced by N_2 gas was supplied for 5 min as the hypoxic respiratory challenge. The hypoxic respiratory challenge was used to mimic breath-holding. After the hypoxic respiratory challenge, 50% O_2 gas balanced by N_2 gas was supplied again for 15 min. The gas modulation was performed using a gas mixer. During the measurement, an automatic temperature controller with a rectal probe (Temperature controller, RWD Life Science, Shenzhen, Guangdong, China) maintained the core temperature of the animal as 37 °C. To ensure the safety of the animal, a pulse oximeter (Mouse Ox, Starr Inc., Oakmont, PA, USA) and a patient monitoring device (B40, GE Healthcare, Chicago, IL, USA) were used to monitor heart rate, oxygen saturation, and the respiratory function of each animal during the measurement.

Following the principles of the 3Rs (replacement, reduction, and refinement) in animal experiments (nc3rs.org.uk (accessed on 2 October 2022)), to minimize the usage of animals, each animal was measured 5 to 6 times on different days. The study was approved by the Institutional Animal Care and Use Committee (IACUC) of the Gwangju Institute of Science and Technology (Protocol number: GIST-2017-065) and the Chonnam National University (protocol number: CNU IACUC-YB-2018-08).

2.5. Modified Beer-Lambert's Law

To extract relative hemoglobin concentration from the DOS raw data, the modified Beer–Lambert law (MBLL) was used [24–26]. Equation (1) shows MBLL with wavelengths of 780, 808, and 820 nm:

$$
\begin{pmatrix} \Delta RHb \\ \Delta OHb \end{pmatrix} = \frac{1}{d \cdot DPF} \begin{pmatrix} \varepsilon_{RHb}^{780\ nm} & \varepsilon_{OHb}^{780\ nm} \\ \varepsilon_{RHb}^{808\ nm} & \varepsilon_{OHb}^{808\ nm} \\ \varepsilon_{RHb}^{820\ nm} & \varepsilon_{OHb}^{820\ nm} \end{pmatrix}^{-1} \begin{pmatrix} \Delta OD^{780\ nm} \\ \Delta OD^{808\ nm} \\ \Delta OD^{820\ nm} \end{pmatrix} \tag{1}
$$

where RHb and OHb denote deoxygenated and oxygenated hemoglobin, respectively, ΔRHb and ΔOHb denote the relative RHb and OHb changes, respectively, d denotes a source –detector separation (7 mm in this study), DPF denotes the differential path length factor, $\varepsilon_{RHb}^{\lambda}$ and $\varepsilon_{OHb}^{\lambda}$ denote the extinction coefficient of RHb and OHb at the different wavelengths, respectively, and ΔOD^{λ} ($=log_{10}\left(I_t^{\lambda}/I_0^{\lambda}\right)$) denotes the transient optical density at the different wavelengths. I_t^{λ} and I_0^{λ} are the transient and baseline intensity at the different wavelengths, respectively. The average intensity of the first 20 s signals at the different wavelengths was considered as I_0^{λ}. The ε_{RHb} and ε_{OHb} at different wavelengths are from a web page hosted by S. Prahl and S. Jacques [27]. The unit of hemoglobin concentration was set to be mM/DPF, which is equivalent to setting DPF as 1 at different wavelengths [24]. Note that considering DPF as a part of the unit of hemoglobin concentration is one of the widely accepted approaches [12,24,25,28]. In this study, 808 nm was selected because it is known as one of the isosbestic points of the extinction coefficients of RHb and OHb [29]. Meanwhile, 780 nm and 820 nm were selected to reflect the different light absorption tendencies between RHb and OHb. Still, other optical wavelengths can be used depending on experimental conditions. Owing to the existence of a non-square matrix of the extinction coefficients in MBLL, the Moore–Penrose pseudoinverse was used to get the inverse of the non-square matrix [26,30].

2.6. Monte Carlo Simulation of Probing Depth

The Monte Carlo simulation is widely used for the simulation of light propagation through various types of media. To validate the probing depth of the DOS system with the 7 mm source-detector separation, a set of Monte Carlo simulations was performed using a MCXLAB toolbox [31]. For simplicity, a three-layer (scalp, skull, and brain), 2D flat slab with a dimension of 10×20 mm^2 (x-axis $\times$ z-axis) was assumed. Table 1 shows the dimensional and optical properties of the scalp, skull, and brain used in the simulation. Due to the unavailability of appropriate absorption coefficients of the scalp, following the work of S. Y. Lee et al. [32], estimated constituents of absorbers of the human extracerebral tissue that treat the scalp and the skull as a single tissue layer [32,33] were used to get absorption coefficients of the scalp at desired wavelengths. In one simulation, the source was set to have a size of 600 μm and a NA of 0.39 to mimic the source fiber in the measurements. In another simulation, the source was set to have a size of 400 μm and a NA of 0.39 to mimic the detector fiber in the measurements. Each simulation was performed by launching 10^7 photons. Two separate simulations performed with the two sources were set to have a 7 mm distance to mimic the 7 mm source-detector separation. Element-wise production was performed on the fluence results of the two simulations to get depth sensitivity maps. Note that this approach is commonly used to get the depth sensitivity map [34,35].

Table 1. Dimensional and optical properties used in Monte Carlo simulation. λ_1: 780 nm; λ_2: 808 nm; λ_3: 820 nm.

	Scalp		Skull		Brain	
Thickness (mm) [36]	0.7		0.6		8.7	
Absorption coefficient (mm^{-1}) [32,33,37]	λ_1	0.0104			λ_1	0.0132
	λ_2	0.0103	0.0250		λ_2	0.0131
	λ_3	0.0107			λ_3	0.0137
Scattering coefficient (mm^{-1}) [18,32,36]	λ_1	9.9982	λ_1	25.2684	λ_1	11.9932
	λ_2	9.7269	λ_2	24.6957	λ_2	11.5491
	λ_3	9.6257	λ_3	24.4602	λ_3	11.3683
Refractive index [32,38–40]	1.38		1.55		1.37	
Anisotropy [32,37,40]	0.80		0.92		0.90	

2.7. Extraction of Hemodynamic Features

To quantify differences between AD and WT, quantitative hemodynamic features were adopted from breast cancer studies that used diffuse optical tomography—an optical method that can acquire 3D absorption and scattering images of biological tissues—and dynamic contrast-enhanced MRI [41,42]. Originally, the features were suggested to quantify the hemodynamic response of the breast during hypoxic respiratory challenge for diagnosing breast cancer and monitoring neoadjuvant chemotherapy efficacy. While we investigated AD in this study because the study also aimed to quantify hemodynamic response caused by hypoxic respiratory challenge, the hemodynamic features used in the breast cancer studies were adopted with a slight modification. Six hemodynamic features derived via simple arithmetic calculation are shown below [41,42]:

$$S_{min} = \min\{S(t)\}; \quad t : 600 - 700 \text{ s} \tag{2}$$

$$S_{max} = max\{S(t)\}; \quad t : 900 - 1000 \text{ s}, \tag{3}$$

$$IE = (S_{max} - S_{min})/S_{max}, \tag{4}$$

$$PIE = (S_{PostBH30sec} - S_{max})/S_{max}, \tag{5}$$

$$m_{rise} = (S_{max} - S_{min})/(T_{max} - T_{min}), \tag{6}$$

$$m_{fall} = (S_{max} - S_{PostBH30sec})/(T_{max} - T_{PostBH30sec}), \tag{7}$$

where $S(t)$ denotes the signal as a function of time, S_{min} denotes the minimum of the signal within 600–700 s, S_{max} denotes the maximum of the signal within 900–1000 s (after the end of hypoxic respiratory challenge), IE denotes initial enhancement, PIE denotes post-initial enhancement, $S_{PostBH30sec}$ denotes the amplitude of the signal at 930 s (30 s after the end of hypoxic respiratory challenge), m_{rise} denotes the slope of the rise within T_{max} and T_{min}, T_{max} and T_{min} denote the corresponding times of S_{max} and S_{min}, respectively, m_{fall} denotes the slope of the drop within T_{max} and $T_{PostBH30sec}$, and $T_{PostBH30sec}$ denotes the corresponding time of $S_{PostBH30sec}$.

The other four parameters are derived via nonlinear fitting of the signal to the equations shown below [41–43]:

$$S(t) = A_{rise,1} \times e^{(q_{rise,1})t} + A_{rise,2} \times e^{(q_{rise,2})t}; \quad t : 600 - 900 \text{ s}, \tag{8}$$

$$S(t) = A_{fall} \times e^{(q_{fall})t}; \quad t : 900 - 1000 \text{ s}, \tag{9}$$

where $A_{rise,\,1}$ and $A_{rise,2}$ denote the amplitudes of the rise of the first and second exponentials, respectively, e denotes the exponential function, $q_{rise,1}$ and $q_{rise,2}$ denote the rate of the rise of the first and second exponentials, respectively, A_{fall} denotes the amplitude of the washout rate, and q_{fall} denotes the rate of the drop. Figure 2 shows the example of the extraction of the hemodynamic features.
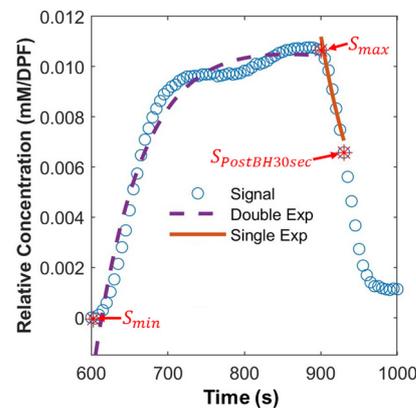
Figure 2. Example of hemodynamic features extraction (from the grand average of ΔRHb of the AD group). Blue dots indicate the signal. S_{min}: the smallest amplitude of the signal between 600 and 700 s; S_{max}: the largest amplitude of the signal between 900 and 1000 s; $S_{PostBH30sec}$: the amplitude at 930 s (30 s after the end of hypoxic respiratory challenge). A purple dashed line indicates the result of the double exponential fitting to the signal between 600 and 900 s. An orange line indicates the result of the single Exponential fitting to the signal between 900 and 930 s.

All of the signals were detrended, up-sampled to 10 Hz, and then smoothed using a moving average filter (window size: 500) before extracting the features described above. In the case of ΔRHb, the features were calculated directly. In the case of ΔOHb, the features were calculated using the negative of the signal ($-\Delta OHb$) due to its trend toward the negative during hypoxic respiratory challenge. Note that the references of the hemodynamic features used signals at 15 s after the end of hypoxic respiratory challenge for calculating some features (e.g., $S_{PostBH15sec}$ and $T_{PostBH15sec}$); however, here signals at 30 s after the end of hypoxic respiratory challenge were used due to the use of anesthesia that may alter the vascular response. For Equation (8), the double exponential was used in this study due to the tendency of the resultant signals while the references used the single exponential.

The features in Equations (8) and (9) (A_{rise1}, A_{rise2}, q_{rise1}, q_{rise2}, A_{fall}, and q_{fall}) were calculated using *fminsearch*, a MATLAB built-in function for nonlinear fitting (MATLAB R2019b, MathWorks, Natick, MA, USA). Among the whole data set, time-series data with issues, including system issues, unstable measurements, and different experimental conditions, were totally excluded and not used throughout the study.

2.8. Statistical Analysis

Before performing the statistical analysis, normality tests were performed for each hemodynamic parameter using the Jarque–Bera test [44]. When a hemodynamic parameter of both AD and WT passed the normality test, the two-sample *t*-test [45] was used to perform the statistical analysis. Otherwise, the Wilcoxon rank-sum test [46] was used to perform the statistical analysis. In this work, a *p*-value lower than 0.05 was considered to be statistically significant. Normality tests and statistical analysis were performed using built-in functions in MATLAB (MATLAB R2019b, MathWorks, Natick, MA, USA).

2.9. Machine Learning (ML)-Based Classification

ML is a widely utilized technique for regression and classification. To further demonstrate the feasibility of the suggested method for differentiating AD and WT, we performed ML-based classification. Among various ML toolboxes, PyCaret was used because PyCaret

is a handy library that offers multiple widely used ML algorithms [47]. ML algorithms used in the study include logistic regression [48], ridge classifier [49], linear discriminant analysis [50,51], K-nearest neighbor classifier [49], support vector machine [52], naive Bayes [49], decision tree classifier [53], gradient boosting algorithm [54], light gradient boosting machine [54], random forest classifier [55], quadratic discriminant analysis [51], extreme gradient boosting [54], AdaBoost classifier [56], extra trees classifier [57], and CatBoost classifier [54]. Note that various ML algorithms have been tested since there is no fixed guidance or solution that determines which algorithm works best for a specific problem. A total of 80% of the data were used as training data to fine-tune each algorithm and 20% of data were used to evaluate the fine-tuned algorithm. K-fold with a fold number of 32 was used due to the limited datasets of the study. For each fold, data were shuffled. Data imbalance was mitigated using a synthetic minority over-sampling technique (SMOTE) [47,58]. Outliers, which were replaced as a not-a-number (NaN) during the extraction of the hemodynamic parameters in Section 2.7, were imputed using a median value of each parameter. Parameters with low variance were ignored during the fine-tuning of each algorithm. Data of each parameter were scaled based on the interquartile range.

To evaluate the performance of each ML algorithm, *accuracy*, *precision*, *recall*, and *F1–score* were used. *Accuracy*, *precision*, *recall*, and *F1–score* are derived by equations shown below [47,59]:

$$Accuracy = (TP + TN)/(TP + TN + FP + FN), \tag{10}$$

$$Precision = TP/(TP + FP), \tag{11}$$

$$Recall = TP/(TP + FN), \tag{12}$$

$$F1 - score = 2\{(Recall \times Precision)/(Recall + Precision)\}, \tag{13}$$

where TP denotes true positive (in case the ML model classifies an input as positive and the input is positive), TN denotes true negative (in case the ML model classifies an input as negative and the input is negative), FP denotes false positive (in case the ML model classifies an input as positive but the input is negative), and FN denotes false negative (in case the ML model classifies an input as negative but the input is positive).

Accuracy indicates how much an ML model correctly classifies the input. *Precision* and *Recall* trade off based on their equations. The numerator of *Precision* and *Recall* is TP while the denominators of *Precision* and *Recall* are $TP + FP$ (all items classified as positive) and $TP + FN$ (all positive items), respectively. $F1 - score$, which considers both *Precision* and *Recall* simultaneously, is a harmonic mean of *Precision* and *Recall*. Each ML algorithm was run 600 times by randomly partitioning training and hold-out data (80% and 20%) to partially mitigate the issue of the limited datasets. Then, the average and 95% confidence intervals of *accuracy*, *precision*, *recall*, and *F1–score* were calculated.

3. Results

3.1. Behavioral Tests

As shown in Figure 3a, the percentage of preference index for the novel object in the 5xFAD group was significantly lower than that of WT mice. The data show that object recognition memory was impaired in 5xFAD mice compared to WT mice. The results indicate that 5xFAD animals are impaired in their spatial learning. As shown in Figure 3b, 5xFAD mice had a significantly longer latency time to escape the platform from day two of the consecutive four days compared to WT mice. This observation indicates that 5xFAD animals are impaired in their memory. Collectively with the results of the novel object indication test, we confirm that the 5xFAD mice used in this work are impaired in terms of spatial learning and recognition memory, which are typical symptoms of AD.
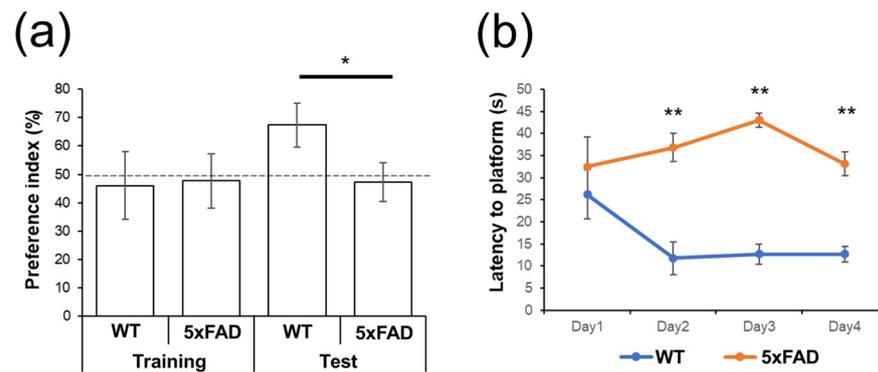
Figure 3. (**a**) Result of novel object recognition test in training and test phases; (**b**) result of Morris water maze as the day passes. All data show means ± standard error of the mean. The significance of differences between the groups was determined using the paired Student's *t*-test. All values of *p* less than 0.05 and 0.005 were considered to be statistically significant and indicated using '*' and '**', respectively.

3.2. Monte Carlo Simulation of Probing Depth

Figure 4 shows the results of Monte Carlo simulations. For better representation of the probing depth and region of the brain, depth sensitivity maps (Figure 4a–c) are superimposed on the mouse brain image that is based on a stereotaxic coordinate image of the mouse brain [60,61]. Due to minute changes in the optical properties of the tissues at different wavelengths, depth sensitivity maps at different wavelengths did not differ significantly. Similarly, normalized sensitivity along the depth did not change much as the wavelength varied (Figure 4d). At the interface between the skull and brain, depth sensitivity at 780 nm, 808 nm, and 820 nm remained to be 0.8357, 0.8332, and 0.8320, respectively. The maximum of depth sensitivity, which existed at the depth of 1.35 mm, of 780 nm, 808 nm, and 820 nm was 0.8497, 0.8448, and 0.8477, respectively. The results of the Monte Carlo simulations showed that more than 83% of photons could reach a depth deeper than 1.3 mm. In the meantime, more than 71% of photons could reach a 3.5 mm depth, which was the expected depth based on the diffusion theory. The sensitivity maps show that the light is mostly probing the cerebral cortex, thus the feasibility of non-invasive brain signal acquisition using the current setup is confirmed. Furthermore, due to the use of close wavelengths, the probing regions of different wavelengths are well overlapped.
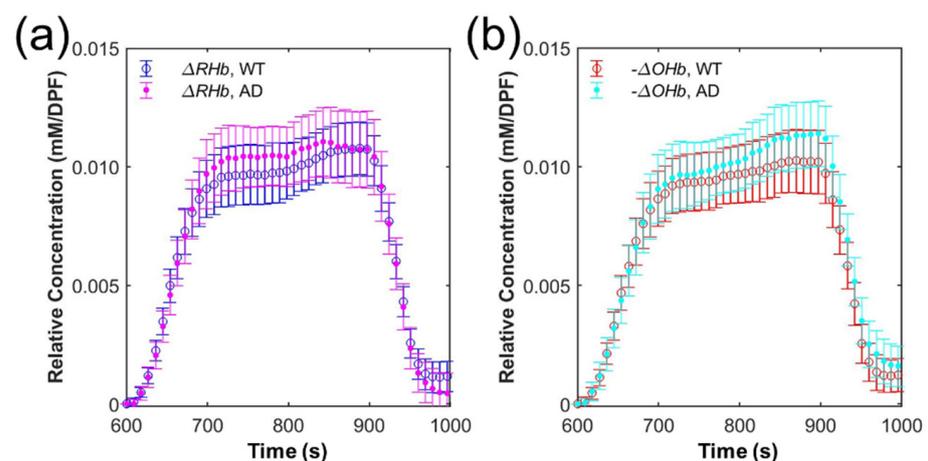
3.3. Grand Average of Hemoglobin Concentration

Figure 5 shows the grand average of ΔRHb and $-\Delta OHb$ during hypoxic respiratory challenge. Hypoxic respiratory challenge was started at 600 s and ended at 900 s. In AD and WT, both ΔRHb and $-\Delta OHb$ showed a rapid increase until approximately 700 s (ΔRHb, AD: 0→0.00971 mM/DPF and WT: 0→0.00904 mM/DPF; $-\Delta OHb$, AD: 0→0.00907 and WT: 0→0.00862) before showing a relatively slow increase or plateau until 900 s (ΔRHb, AD: 0.00971 mM/DPF →0.01072 mM/DPF and WT: 0.00904 mM/DPF→0.01065 mM/DPF; $-\Delta OHb$, AD: 0.00907 mM/DPF→0.01138 mM/DPF and WT: 0.00862 mM/DPF→0.01012 mM/DPF). The increase of ΔRHb and $-\Delta OHb$ had a tendency of the double Exponential function (e.g., $A_1 e^{r_1 t} + A_2 e^{r_2 t}$). After the end of hypoxic respiratory challenge at 900 s, all of ΔRHb and $-\Delta OHb$ promptly decreased until approximately 960 s.

While ΔRHb and $-\Delta OHb$ showed similar trends in both WT and AD, they had some minute differences. ΔRHb of WT reached its maximum (0.01079 mM/DPF) at 887.1 s while the one of AD reached its maximum (0.01112 mM/DPF) at 845.3 s. $-\Delta OHb$ of WT reached its maximum (0.01030 mM/DPF, minimum of ΔOHb) at 866.3 s while the one of AD reached its maximum (0.01143 mM/DPF) at 891.2 s. Note that we analyzed $-\Delta OHb$, not ΔOHb, in order to be consistent with the use of $-\Delta OHb$ in the analysis of the hemodynamic features.
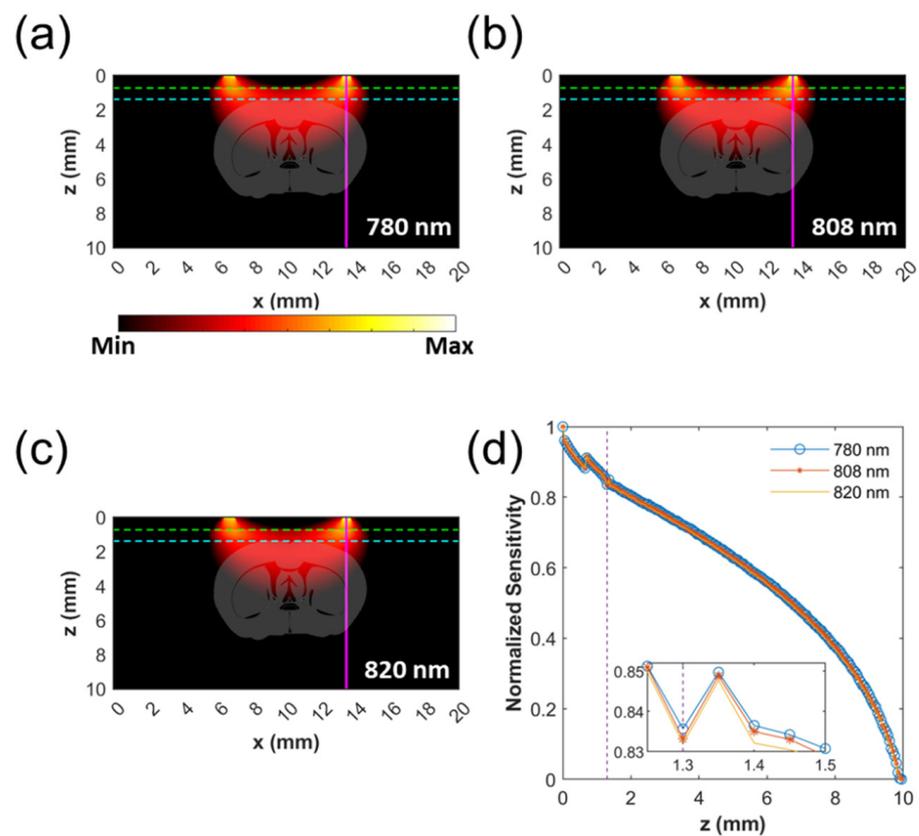
Figure 4. Depth sensitivity maps at wavelengths of 780 nm (**a**), 808 nm (**b**), and 820 nm (**c**) from the Monte Carlo simulation. In (**a**–**c**), green and cyan dotted lines represent boundaries between the scalp and the skull, and the skull and the brain, respectively. The probing depth is deep enough to reach parts of the cerebral cortex. (**d**) Normalized depth sensitivity through a magenta line of (**a**–**c**). A Purple dotted line represents a boundary (1.3 mm) between the skull and the brain. Inset shows a magnified image of normalized sensitivity at depths between 1.25 mm and 1.5 mm.

Figure 5. (**a**) Grand average and 95% confidence interval of relative deoxyhemoglobin concentration change (ΔRHb) of wild-type (WT) (blue circle) and Alzheimer's disease (AD) (purple dot) mice. (**b**) Grand average and 95% confidence interval of the negative relative oxyhemoglobin concentration change ($-\Delta OHb$) of WT (red circle) and AD (cyan dot).

3.4. Statistical Analysis

Tables 2 and 3 show summaries of 13 hemodynamic features and their statistical analyses of ΔRHb and $-\Delta OHb$, respectively. In ΔRHb, $A_{rise,2}$ and $q_{rise,2}$ showed p-values of 0.0086 and 0.0066, respectively. While A_{fall} and $q_{rise,1}$ of ΔRHb showed quite a large difference in the mean between AD and WT, their p-values were greater than 0.05 (0.2798 and 0.8052). In $-\Delta OHb$, m_{rise} showed a p-value of 0.0499. Concurrently, $A_{rise,1}$, $q_{rise,1}$, and S_{max} of $-\Delta OHb$ showed quite a large difference in the mean between AD and WT, even though their p-values were greater than 0.05 (0.2485, 0.1338, and 0.0566).

Table 2. Summary of 13 hemodynamic features of ΔRHb from WT and AD. std: standard deviation. p-value lower than 0.05 is highlighted using red color.

Features	WT (Mean ± std)	AD (Mean ± std)	p-Value	Both Passed Normality Test?
IE	$1.03 \times 10^0 \pm 1.68 \times 10^{-2}$	$1.03 \times 10^0 \pm 3.53 \times 10^{-2}$	0.8429	×
PIE	$-4.00 \times 10^{-1} \pm 6.61 \times 10^{-2}$	$-3.94 \times 10^{-1} \pm 1.10 \times 10^{-1}$	0.7950	○
m_{rise}	$3.63 \times 10^{-5} \pm 5.45 \times 10^{-6}$	$3.79 \times 10^{-5} \pm 1.46 \times 10^{-5}$	0.6190	○
m_{fall}	$-1.43 \times 10^{-4} \pm 4.29 \times 10^{-5}$	$-1.61 \times 10^{-4} \pm 6.47 \times 10^{-5}$	0.2389	○
A_{fall}	$6.31 \times 10^4 \pm 8.89 \times 10^4$	$7.21 \times 10^5 \pm 1.88 \times 10^6$	0.2798	×
$A_{rise,1}$	$2.26 \times 10^{-2} \pm 3.70 \times 10^{-2}$	$3.81 \times 10^{-2} \pm 6.45 \times 10^{-2}$	0.9298	×
$A_{rise,2}$	$-4.70 \times 10^2 \pm 9.70 \times 10^2$	$-4.71 \times 10^1 \pm 6.67 \times 10^1$	0.0086	×
q_{fall}	$1.58 \times 10^{-2} \pm 4.09 \times 10^{-3}$	$1.79 \times 10^{-2} \pm 5.23 \times 10^{-3}$	0.1143	○
$q_{rise,1}$	$-1.06 \times 10^{-3} \pm 2.84 \times 10^{-3}$	$-8.34 \times 10^{-4} \pm 3.77 \times 10^{-3}$	0.8052	○
$q_{rise,2}$	$-1.62 \times 10^{-2} \pm 8.82 \times 10^{-3}$	$-9.68 \times 10^{-3} \pm 7.23 \times 10^{-3}$	0.0066	○
S_{max}	$1.02 \times 10^{-2} \pm 1.87 \times 10^{-3}$	$-1.09 \times 10^{-2} \pm 4.51 \times 10^{-3}$	0.4879	○
S_{min}	$-2.67 \times 10^{-4} \pm 1.81 \times 10^{-4}$	$-2.98 \times 10^{-4} \pm 2.60 \times 10^{-4}$	0.6340	○
$S_{PostBH30sec}$	$6.00 \times 10^{-3} \pm 1.23 \times 10^{-3}$	$6.15 \times 10^{-3} \pm 2.56 \times 10^{-3}$	0.8050	○

Table 3. Summary of 13 hemodynamic features of $-\Delta OHb$ from WT and AD. std: standard deviation. p-value lower than 0.05 is highlighted using red color.

Features	WT (Mean ± std)	AD (Mean ± std)	p-Value	Both Passed Normality Test?
IE	$1.03 \times 10^0 \pm 2.48 \times 10^{-2}$	$1.03 \times 10^0 \pm 2.30 \times 10^{-2}$	0.3207	×
PIE	$-3.92 \times 10^{-1} \pm 1.10 \times 10^{-1}$	$-3.61 \times 10^{-1} \pm 1.34 \times 10^{-1}$	0.3518	○
m_{rise}	$3.44 \times 10^{-5} \pm 8.60 \times 10^{-6}$	$4.04 \times 10^{-5} \pm 1.26 \times 10^{-5}$	0.0499	○
m_{fall}	$-1.33 \times 10^{-4} \pm 4.14 \times 10^{-5}$	$-1.56 \times 10^{-4} \pm 6.14 \times 10^{-5}$	0.1172	○
A_{fall}	$2.33 \times 10^5 \pm 4.16 \times 10^5$	$3.35 \times 10^5 \pm 1.02 \times 10^6$	0.4202	×
$A_{rise,1}$	$2.09 \times 10^0 \pm 4.61 \times 10^0$	$1.33 \times 10^{-2} \pm 1.28 \times 10^{-2}$	0.2485	×
$A_{rise,2}$	$-3.56 \times 10^2 \pm 5.98 \times 10^2$	$-4.47 \times 10^2 \pm 8.82 \times 10^2$	0.8004	×
q_{fall}	$1.65 \times 10^{-2} \pm 5.59 \times 10^{-3}$	$1.69 \times 10^{-2} \pm 6.79 \times 10^{-3}$	0.8486	○
$q_{rise,1}$	$-1.88 \times 10^{-3} \pm 3.31 \times 10^{-3}$	$-6.45 \times 10^{-4} \pm 2.69 \times 10^{-3}$	0.1338	○
$q_{rise,2}$	$-1.37 \times 10^{-2} \pm 6.79 \times 10^{-3}$	$-1.46 \times 10^{-2} \pm 7.35 \times 10^{-3}$	0.6446	○
S_{max}	$9.81 \times 10^{-3} \pm 2.61 \times 10^{-3}$	$1.16 \times 10^{-2} \pm 3.83 \times 10^{-3}$	0.0566	○
S_{min}	$-2.68 \times 10^{-4} \pm 2.45 \times 10^{-4}$	$-3.29 \times 10^{-4} \pm 2.38 \times 10^{-4}$	0.3751	○
$S_{PostBH30sec}$	$6.07 \times 10^{-3} \pm 2.31 \times 10^{-3}$	$7.49 \times 10^{-3} \pm 3.04 \times 10^{-3}$	0.0615	○

3.5. Machine Learning (ML)-Based Classification

The ML-based classification was performed on three different datasets: data with p-values lower than 0.05 ($A_{rise,2}$ and $q_{rise,2}$ of ΔRHb, and m_{rise} of $-\Delta OHb$), whole hemodynamic features of ΔRHb and $-\Delta OHb$ (23 hemodynamic features in total), and three principal components (PCs) from the whole hemodynamic features extracted via principal component analysis (PCA). PCA is a technique that is commonly used for the dimension reduction of multi-dimensional data to extract PCs with large variances [47,62]. The first PC is the extracted data with the largest variance, and the contribution of PCs to the variation of the data decreases as the order of PC increases (e.g., PC1 has a larger variance than PC2, PC2 has a larger variance than PC3, and so forth). Table 4 shows the classification results using data with p-values lower than 0.05. When data with p-values lower than 0.05 were used, logistic regression performed the best showing the best *accuracy* of 62.9%, followed by naive Bayes and quadratic discriminant analysis with an *accuracy* of 62.4%. Table 5 shows the classification results from the whole hemodynamic features. When the whole features were used, naive Bayes performed the best and showed the best *accuracy* of 84.3%. Table 6 shows the classification results using three PCs extracted from the whole hemodynamic features. When three PCs from the whole data were used, the extra trees classifier performed the best and showed the best *accuracy* of 76%. Among the best performing classifiers of different data sets (data with low p-values, whole hemodynamic features, and three PCs derived from the whole hemodynamic features), Naive Bayes using the whole hemodynamic feature performed the best in terms of *accuracy*. Logistic regression using data with low p-values performed the worst.

Table 4. Averaged *accuracy*, *recall*, *precision*, and $F1 - score$ and their 95% confidence interval (CI) of 15 machine learning (ML) methods using hemodynamic features with p-values lower than 0.05 ($A_{rise,2}$ and $q_{rise,2}$ of ΔRHb, and m_{rise} of $-\Delta OHb$) for classification of AD and WT. Each method was performed 600 times by shuffling the data due to limited data sets. The results and the name of the ML method with the best *accuracy* (logistic regression) are highlighted in red.

Method	*Accuracy* (%) [95% CI]	*Recall* (%) [95% CI]	*Precision* (%) [95% CI]	*F1–Score* (%) [95% CI]
Logistic regression	62.9 [62.0, 63.9]	35.8 [34.1, 37.4]	71.8 [69.1, 74.5]	45.1 [43.5, 46.8]
Ridge classifier	52.4 [51.2, 53.5]	30.0 [28.5, 31.5]	62.2 [59.8, 64.6]	35.9 [31.7, 37.1]
Linear discriminant analysis	53.8 [52.6, 55.1]	7.3 [6.4, 8.1]	37.6 [33.7, 41.5]	11.9 [10.6, 13.2]
K-nearest neighbor classifier	51.9 [51.1, 52.7]	18.3 [16.9, 19.8]	42.1 [38.8, 45.3]	22.3 [20.7, 23.9]
Support vector machine	57.2 [56.2, 58.2]	52.2 [50.0, 54.4]	55.1 [52.7, 57.4]	48.4 [46.4, 50.4]
Naive Bayes	62.4 [61.2, 63.7]	25.7 [24.5, 26.9]	84.2 [81.3, 87.1]	38.4 [36.8, 40.0]
Decision tree classifier	50.1 [48.9, 51.4]	0.6 [0.3, 1.0]	1.5 [0.6, 2.3]	0.8 [0.3, 1.2]
Gradient boosting classifier	46.0 [44.9, 47.2]	60.8 [58.4, 63.2]	51.8 [50.0, 53.6]	48.6 [47.4, 49.9]
Light gradient boosting machine	55.1 [54.2, 55.9]	12.7 [11.0, 14.5]	18.0 [15.5, 20.5]	14.7 [12.7, 16.7]
Random forest classifier	44.6 [43.7, 45.5]	41.1 [37.6, 44.6]	32.6 [29.9, 35.4]	28.1 [26.0, 30.3]

Table 4. *Cont.*

Method	Accuracy (%) [95% CI]	Recall (%) [95% CI]	Precision (%) [95% CI]	F1−Score (%) [95% CI]
Quadratic discriminant analysis	62.4 [61.4, 63.4]	27.5 [26.3, 28.6]	79.0 [76.1, 81.9]	39.7 [38.2, 41.2]
Extreme gradient boosting	49.5 [48.4, 50.7]	63.0 [59.3, 66.8]	32.3 [30.1, 34.5]	41.3 [38.7, 43.9]
AdaBoost classifier	59.5 [58.2, 60.8]	24.7 [23.2, 26.2]	67.8 [64.3, 71.2]	34.4 [32.5, 36.4]
Extra trees classifier	38.6 [37.6, 39.6]	39.1 [35.5, 42.8]	16.2 [14.6, 17.7]	22.3 [20.2, 24.4]
CatBoost classifier	46.7 [45.2, 48.2]	26.7 [24.8, 28.6]	44.8 [41.8, 47.8]	28.5 [26.9, 30.1]

Table 5. Averaged *accuracy*, *recall*, *precision*, and $F1 - score$ and their 95% CI of 15 ML methods using hemodynamic features with whole hemodynamic features for classification of AD and WT. Each method was performed 600 times by shuffling the data due to limited data sets. The results and the name of the ML method with the best *accuracy* (naive Bayes) are highlighted in red.

Method	Accuracy (%) [95% CI]	Recall (%) [95% CI]	Precision (%) [95% CI]	F1−Score (%) [95% CI]
Logistic regression	71.8 [70.8, 72.8]	69.7 [68.3, 71.0]	65.3 [63.7, 66.9]	65.1 [64.0, 66.2]
Ridge classifier	49.7 [48.7, 50.6]	42.0 [40.3, 43.8]	42.9 [40.9, 44.9]	40.3 [38.7, 41.9]
Linear discriminant analysis	67.0 [65.8, 68.2]	28.0 [26.0, 30.0]	64.3 [60.6, 67.9]	34.0 [31.9, 36.1]
K-nearest neighbor classifier	75.1 [74.0, 76.1]	43.5 [41.3, 45.7]	72.4 [69.7, 75.1]	51.9 [49.6, 54.2]
Support vector machine	66.7 [65.5, 67.8]	69.3 [66.8, 71.7]	50.4 [48.3, 52.5]	56.9 [54.8, 59.0]
Naive Bayes	84.3 [83.8, 84.8]	69.7 [68.4, 71.1]	89.9 [88.4, 91.4]	75.4 [74.4, 76.3]
Decision tree classifier	52.7 [51.2, 54.2]	24.6 [22.7, 26.5]	25.6 [23.5, 27.7]	23.4 [21.6, 25.1]
Gradient boosting classifier	51.6 [50.1, 53.1]	52.7 [50.2, 55.1]	41.5 [39.4, 43.6]	42.8 [40.9, 44.7]
Light gradient boosting machine	54.2 [52.5, 56.0]	16.7 [14.1, 19.4]	13.9 [11.9, 15.9]	14.2 [12.1, 16.3]
Random forest classifier	54.0 [52.5, 55.5]	28.8 [26.2, 31.5]	27.3 [25.1, 29.5]	24.8 [22.8, 26.8]
Quadratic discriminant analysis	81.3 [80.6, 81.9]	75.4 [74.0, 76.8]	76.6 [75.3, 77.8]	74.0 [73.0, 75.0]
Extreme gradient boosting	39.5 [38.4, 40.5]	78.9 [76.1, 81.6]	34.6 [33.1, 36.1]	44.9 [43.1, 46.7]
AdaBoost classifier	59.1 [58.1, 60.2]	39.5 [37.1, 41.8]	43.2 [40.5, 45.9]	39.6 [37.3, 42.0]
Extra trees classifier	54.4 [53.0, 55.8]	19.7 [16.9, 22.5]	17.8 [15.3, 20.4]	13.2 [11.5, 14.9]
CatBoost classifier	49.1 [47.5, 50.7]	38.3 [36.0, 40.6]	39.2 [36.8, 41.6]	38.1 [35.8, 40.3]

Table 6. Averaged *accuracy*, *recall*, *precision*, and *F*1 − *score* and their 95% CI of 15 ML methods using three principal components (PCs) of principal component analysis (PCA) extracted from whole hemodynamic features for classification of AD and WT. Each method was performed 600 times by shuffling the data due to limited data sets. The results and the name of the ML method with the best *accuracy* (extra trees classifier) are highlighted in red.

Method	*Accuracy* (%) [95% CI]	*Recall* (%) [95% CI]	*Precision* (%) [95% CI]	*F*1–*Score* (%) [95% CI]
Logistic regression	73.8 [72.9, 74.7]	83.2 [82.0, 84.4]	69.3 [68.0, 70.5]	74.1 [73.1, 75.0]
Ridge classifier	73.5 [72.6, 74.4]	83.2 [82.0, 84.4]	68.9 [67.6, 70.2]	73.9 [73.0, 74.8]
Linear discriminant analysis	73.7 [72.8, 74.6]	83.2 [82.0, 84.4]	69.1 [67.9, 70.4]	74.0 [73.1, 74.9]
K-nearest neighbor classifier	73.8 [72.8, 74.7]	80.9 [79.4, 82.5]	68.9 [67.5, 70.4]	72.9 [71.6, 74.1]
Support vector machine	71.1 [70.3, 71.9]	81.7 [80.2, 83.1]	65.5 [64.2, 66.9]	71.1 [69.9, 72.2]
Naive Bayes	73.3 [72.5, 74.1]	85.6 [84.4, 86.7]	67.8 [66.6, 69.0]	74.1 [73.3, 75.0]
Decision tree classifier	61.5 [60.2, 62.7]	37.3 [34.3, 40.3]	47.2 [44.1, 50.2]	37.8 [35.1, 40.5]
Gradient boosting classifier	72.4 [71.2, 73.5]	67.4 [64.8, 70.0]	69.1 [67.0, 71.2]	64.5 [62.4, 66.7]
Light gradient boosting machine	65.7 [64.3, 67.0]	41.2 [37.8, 44.5]	41.9 [38.7, 45.0]	39.2 [36.2, 42.3]
Random forest classifier	69.0 [67.7, 70.2]	58.3 [55.5, 61.2]	62.9 [60.5, 65.4]	56.3 [53.9, 58.7]
Quadratic discriminant analysis	72.7 [718, 73.6]	83.8 [82.6, 85.0]	67.6 [66.4, 68.9]	73.4 [72.5, 74.3]
Extreme gradient boosting	63.1 [62.0, 64.1]	73.6 [71.4, 75.8]	60.1 [58.4, 61.7]	62.0 [60.5, 63.6]
AdaBoost classifier	65.2 [63.9, 66.5]	49.1 [46.1, 52.1]	56.2 [53.3, 59.1]	47.9 [45.3, 50.5]
Extra trees classifier	76.0 [75.1, 76.9]	80.1 [78.5, 81.7]	72.3 [70.9, 73.7]	74.2 [73.0, 75.4]
CatBoost classifier	65.8 [64.6, 67.0]	53.4 [50.5, 56.3]	60.5 [57.9, 63.1]	57.1 [49.3, 54.1]

4. Discussion

In this work, the hypoxic respiratory challenge was used to induce hemodynamic reactivity in an animal model. Since animals cannot perform breath-holding spontaneously, the hypoxic respiratory challenge was used to mimic breath-holding even though hypoxic respiratory challenge may not be able to mimic breath-holding perfectly. The experimental protocol used in the work (50% O_2 gas balanced by N_2 gas for 10 min → 16% O_2 gas balanced by N_2 gas for 5 min as hypoxic gas challenge → 50% O_2 gas balanced by N_2 gas for 15 min) was based on a protocol used in a previous breast cancer study by S. Lee and J. G. Kim (100% O_2 gas for 3 min → 21% O_2 gas balanced by N_2 gas for 10 min as the hypoxic gas challenge) with a minute modification [28]. In the previous study, the average pulmonary oxygen saturation (SpO_2) of rats changed from 97% to 91% as the supplied gas was changed from 100% O_2 gas to 21% O_2 gas balanced by N_2 gas. In this study, the average SpO_2 changed from 98.6% to 94.4% as the supplied gas was changed from 50% O_2 gas balanced by N_2 gas to 16% O_2 gas balanced by N_2 gas. While the drop of SpO_2 in this study was smaller (98.6% → 94.4%: 4.2% difference) than the one of the previous study (97% → 91%: 6% difference) [28], the drop of SpO_2, which is observed in a human study during breath-holding [63], guarantees that the protocol used in this study can mimic

breath-holding as well. At the same time, 94.4% SpO_2 can assure the safety of the animal because this value is higher than the one (92% SpO_2) of people living in Aspen, Colorado where people do not have any critical health issues [28]. The discrepancy between the study of S. Lee and J. G. Kim and this study may come from the use of different experimental protocols and the difference in animals used. While the protocol—either by letting a subject perform breath-holding spontaneously or supplying hypoxic gas—can be relatively easily carried out, the use of the protocol should be avoided for people with respiratory diseases.

As the ultimate goal of the suggested technique is to translate the technique to clinics in the future, the total amount of time for performing signal acquisition and processing—including the conversion of the raw data to hemoglobin concentration data and classification using an ML algorithm—should be short enough. In the animal experiments, assuming that all preparations—including shaving and depilating the fur on the leg and the head—are done, it takes 30 min for signal acquisition. The signal acquisition time can be shortened to several minutes maximum, which would include minutes of baseline and recovery measurements and breath holding or breathing of hypoxic gas if the signal acquisition is performed on human subjects. The conversion of optical intensity data to hemoglobin concentration and the extraction of hemodynamic features takes less than one minute, and may take a shorter length of time if the amount of the data becomes smaller (e.g., via the use of a shorter experimental protocol). For classification, it would take several seconds at maximum using a pre-trained ML algorithm. Overall, the total time may take less than 40 min for animals in this work and several minutes for humans. Thus, the signal acquisition and processing time of the technique will not be a hurdle when translating the technique to the clinics in the future.

Other than this work, there have been studies that attempted using breath-holding in mild cognitive impairment (MCI) and AD studies. Rather than attempting to extract various features from CVR signals, most of the previous MCI and AD studies that utilized transcranial Doppler for monitoring blood flow used the so-called breath-holding index, which is relatively simple [64,65]. Even though this work is not the first work that uses breath-holding for investigating AD or MCI, by using the simple DOS system for the acquisition of hemodynamic signals, extracting hemodynamic features from CVR caused by breath-holding, and adopting modern ML algorithms, this work attempted to examine an alternative approach for differentiating AD and WT. Thus, we believe that such a new approach will be able to encourage researchers to investigate the use of hemodynamic features from CVR signals caused by breath-holding or other types of breathing gas (e.g., carbogen) for other brain diseases that cause the impairment of the cerebrovasculature. In the meantime, changes in hemodynamic features have also been studied in animal studies to test cyanide toxicity [26,66], breast cancer studies [41], and other brain diseases such as Parkinson's disease [67] and traumatic brain injury [68]. We believe that the suggested approach, with modifications, has the potential to be used in cases other than AD, such as the cases mentioned above.

In this work, the best ML result (84% *accuracy*) was achieved when the whole hemodynamic features were used as input features, and naive Bayes was employed for binary classification of AD and WT. As mentioned in the introduction, we note that some previous studies attempted to use AI, including ML and DL, for the classification of a sub-type of AD based on hemodynamic signals provoked by various cognitive tests. A study by D. Yang et al. adopted N-back, Stroop, and verbal fluency tasks as interventions to classify MCI patients and healthy controls [10]. A commercial, multi-channel functional near-infrared spectroscopy (fNIRS) was utilized to acquire hemoglobin concentration maps in the study. Using a convolutional neural network (CNN)-based AI algorithm, the researchers achieved 90% *accuracy* using the verbal fluency test, and achieved the best *accuracy* of 98.61% using the N-back test. Another study by D. Yang et al. attempted to use resting-state fNIRS signals acquired using the same commercial fNIRS system for the classification of MCI patients and healthy controls [69]. A connectivity map of ΔOHb and ΔRHb features extracted using various pre-trained CNN models (e.g., variations of VGG, Densenet, Alexnet,

and Resnet networks) for transfer learning, time-series of ΔOHb, and ΔRHb were used as input features for ML classifiers, including linear discriminant analysis, support vector machine, and K-nearest neighbor. As the time-series signals were given as inputs, the best classification *accuracy* only reached 67% using linear discriminant analysis. Interestingly, because the features extracted using VGG19, a type of widely used CNN model, were used, the best *accuracy* of 95.81% could be achieved when the support vector machine was used as a classifier. In the meantime, one study by T. K. K. Ho et al. adopted oddball, 1-back, and verbal fluency tasks as interventions and attempted to perform classification of multi-stage AD with the inclusion of healthy controls, asymptomatic AD (a group of people with amyloid deposition but without cognitive impairment), prodromal AD (a group of people with MCI), and AD (a group of people with significant cognitive impairment due to AD) [12]. A lab-built, multi-channel fNIRS was utilized to acquire time-series hemoglobin concentration signals in the study. For classification, the researchers tested various ML (linear discriminant analysis, k-nearest neighbor, Gaussian naive Bayes, support vector machine, AdaBoost, random forest, and ensemble learning) and DL (simple neural network, 1D-CNN, long-short term memory (LSTM), gated recurrent units, and a combination of CNN and LSTM (CNN-LSTM)) algorithms. Among ML algorithms, using ΔOHb, ensemble learning achieved the best *accuracy* of 82.9%. Among DL algorithms, using relative total hemoglobin concentration (ΔTHb), which is calculated via summation of ΔOHb and ΔRHb (i.e., $\Delta THb = \Delta OHb + \Delta RHb$), CNN-LSTM achieved an *accuracy* of 87.7%. The studies used more advanced AI algorithms as either a classifier or a feature extractor and achieved a better *accuracy* than the current study. However, the best *accuracy* in this study outperformed the one in one study by D. Yang et al. [69] with ML classifiers as ΔOHb and ΔRHb were used as inputs (84% vs. 67%), even though the direct comparison between the study of D. Yang et al. and this work is not straightforward due to a few differences such as differences in subjects. Still, by comparing results from pure ML classifiers and a combination of AI algorithms, the study of D. Yang et al. [69] shows the potential of using advanced AI techniques to fully extract hidden information in hemodynamic signals. In this work, we could not test more advanced AI techniques due to the use of a single-channel system and limited datasets. A limited population is another factor that limited the use of more advanced AI techniques. In the future, a study using a multi-channel system and a larger population will need to be performed to more rigorously test the use of breath-holding or hypoxic respiratory challenge in AD.

We postulated that CVR, which was measured using the DOS system, during breath-holding may be able to be used to differentiate AD and WT due to differences in the cerebrovasculature. While the time for onset and symptoms vary depending on AD animal models, many AD models show cerebrovascular abnormalities [21,22,70,71]. In the 5xFAD model, as β-amyloid accumulates, morphological abnormality of the cerebrovasculature, dysfunction of the blood–brain barrier, and variations in cerebrovascular density happen, which may be caused by or progressed due to inflammatory response and cerebral amyloid angiopathy [20–23]. In turn, neurovascular units (NVUs), which are responsible for maintaining homeostasis of the cerebrovasculature, are deteriorated and may fail their functions. As the deterioration of NVUs and the cerebrovasculature also happens in the 5xFAD model, we could observe CVR differences between AD and WT. ΔRHb in AD showed a larger maximum and took a shorter time to reach the maximum than ΔRHb in WT (AD: 0.01112 mM/DPF (845.3 s) and WT: 0.01079 mM/DPF (887.1 s)). Meanwhile, ΔOHb in AD showed a smaller minimum and took a longer time to reach the minimum than ΔOHb in WT (AD: -0.01143 mM/DPF (891.2 s) and WT: -0.01030 mM/DPF (866.3 s)). Additionally, as shown in Figure 6, starting from approximately 820 s, even after the end of hypoxic respiratory challenge at 900 s, ΔTHb in AD kept decreasing while ΔTHb in WT was maintained around the baseline level. The differences in ΔRHb, ΔOHb, and ΔTHb between AD and WT indicate that cerebrovascular impairment exists in the 5xFAD model. Such differences may originate from the impairments mentioned above, which can lead to the disruption of the regulation of cerebral blood flow and the exchange of O_2 and

CO_2. We believe that these differences could be more quantitatively analyzed by utilizing 13 hemodynamic features. However, ML results using the whole hemodynamic features were better than the ones of other features used in this study. Since similar impairments happen in other AD models and even in clinical subjects [22,23,64,65,72], we expect that our technique can be translated to the clinics with some modifications.
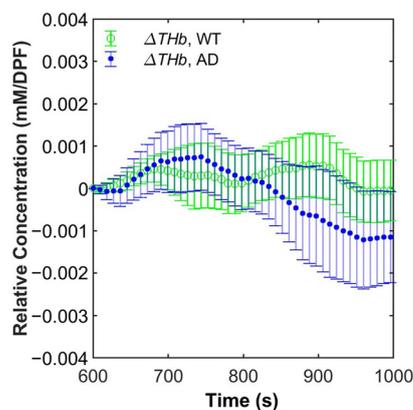
Figure 6. Grand average and 95% confidence interval of relative total hemoglobin concentration (ΔTHb) of WT (green circle) and AD (blue dot).

As described above, this study has some limitations that should be addressed in the future. First, this study used a relatively small population. Even though each animal was measured multiple times to partially address the population issue, this approach may not fully resolve the issue of the limited population. In the meantime, the acquisition of signals from larger populations will be able to allow the use of more advanced AI techniques such as deep learning so that more advanced analyses, including multi-stage classification of AD, can be performed. Thus, a follow-up study with larger populations will need to be performed. Second, the DOS measurement was performed using a single source-detector separation. While the results of the Monte Carlo simulation guarantee the feasibility of the brain signal measurement using the DOS system, based on diffuse optics theories, deep tissue measurements of multi-layered tissues (e.g., the scalp, skull, and brain) are affected by the changes in optical properties of both extracerebral (e.g., the scalp and skull) and intracerebral tissues (e.g., the brain) [32]. To minimize the extracerebral signals, a multi-channel system with multiple source-detector pairs (e.g., with both short separation channels and long separation channels) can be used in the future [73]. Another advantage of using the multi-channel system includes the possibility of the acquisition of signals from multiple brain regions that can allow the use of more advanced analysis (e.g., functional connectivity) and more advanced AI techniques. Additionally, the use of a multi-modal system, including a combination of an optical system and an electroencephalogram [73], can be another interesting work that is worth investigating in the future. Furthermore, cytochrome c oxidase (CCO), which is closely related to the oxidative metabolism of mitochondria and the synthesis of adenosine triphosphate (ATP), can be another potential parameter that can be investigated due to the possibility of using DOS for CCO measurement [74] and the relationship between the decrease of CCO activity and the progress of AD [75,76]. Third, to perform a stable signal measurement, animals were anesthetized during the signal acquisition. While anesthesia is widely utilized in animal studies, the physiological response may be altered by anesthesia. In this study, hypoxic respiratory challenge, which was used to mimic breath-holding, was maintained for a relatively long time (5 min) to compensate the delayed response caused by anesthesia. However, this may not happen in human experiments. Rather, breath-holding is mostly performed for less than 1 min in human experiments. To mimic more realistic breath-holding in animal experiments, the animal can be placed in a container that is much larger than a usual induction chamber of anesthesia and different types of breathing gas can be supplied. In the meantime, a miniature, wireless

system that can be fixed on the head can be used to monitor the brain response during breath-holding. Fourth, the best *accuracy* has not reached the level of similar studies. While the suggested method has the potential to be used in the early diagnosis of AD, there are many spaces where the method should be further improved to use the method on patients due to the simplicity of the method. Lastly, we only used a 5xFAD AD model and WT animals in this work. In the future, the same work may need to be performed with various types of AD models, including not only models of FAD but also models of sporadic AD, which is found to be about 95% of AD cases [22].

5. Conclusions

In conclusion, we performed a preclinical study to test the feasibility of combining hypoxic respiratory challenge, which was used to mimic breath-holding, and diffuse optical signal measurements to differentiate AD and WT. To quantify hemodynamic response during hypoxic respiratory challenge, we employed hemodynamic features that were originally used to study the hemodynamic response of breast cancer during breath-holding. To further confirm the usefulness of our approach, we used various ML techniques to classify AD and WT. Using whole hemodynamic features, the best *accuracy* of 84% was achieved. In the future, we will perform experiments with a larger population using multi-channel brain monitoring systems to investigate the applicability of the breath-holding protocol for classification of multi-stage AD.

Author Contributions: Conceptualization, M.S. and J.G.K.; methodology, M.S. and Y.O.; software, M.S. and Y.O.; validation, M.S.; formal analysis, M.S.; resources, H.J.P. and W.-S.C.; writing—original draft preparation, M.S.; writing—review and editing, M.S., Y.O., J.G.K., H.J.P. and W.-S.C.; supervision, W.-S.C. and J.G.K.; project administration, W.-S.C. and J.G.K.; funding acquisition, M.S., W.-S.C. and J.G.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Animal Care and Use Committee of the Gwangju Institute of Science and Technology (Protocol number: GIST-2017-065 (2017-09)) and the Chonnam National University (CNU IACUC-YB-2018-08 (2018-08)).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

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

References

1. Alzheimer's Association. 2020 Alzheimer's Disease Facts and Figures. *Alzheimers Dement.* **2020**, *16*, 391–460. [CrossRef]
2. Vlassenko, A.G.; Benzinger, T.L.S.; Morris, J.C. PET Amyloid-Beta Imaging in Preclinical Alzheimer's Disease. *Biochim. Biophys. Acta BBA-Mol. Basis Dis.* **2012**, *1822*, 370–379. [CrossRef] [PubMed]
3. Sehlin, D.; Fang, X.T.; Cato, L.; Antoni, G.; Lannfelt, L.; Syvänen, S. Antibody-Based PET Imaging of Amyloid Beta in Mouse Models of Alzheimer's Disease. *Nat. Commun.* **2016**, *7*, 10759. [CrossRef] [PubMed]
4. Wang, L.; Benzinger, T.L.; Su, Y.; Christensen, J.; Friedrichsen, K.; Aldea, P.; McConathy, J.; Cairns, N.J.; Fagan, A.M.; Morris, J.C.; et al. Evaluation of Tau Imaging in Staging Alzheimer Disease and Revealing Interactions between β-Amyloid and Tauopathy. *JAMA Neurol.* **2016**, *73*, 1070. [CrossRef]
5. Cho, H.; Choi, J.Y.; Hwang, M.S.; Lee, J.H.; Kim, Y.J.; Lee, H.M.; Lyoo, C.H.; Ryu, Y.H.; Lee, M.S. Tau PET in Alzheimer Disease and Mild Cognitive Impairment. *Neurology* **2016**, *87*, 375–383. [CrossRef]

6. Cantin, S.; Villien, M.; Moreaud, O.; Tropres, I.; Keignart, S.; Chipon, E.; Le Bas, J.-F.; Warnking, J.; Krainik, A. Impaired Cerebral Vasoreactivity to CO2 in Alzheimer's Disease Using BOLD FMRI. *NeuroImage* **2011**, *58*, 579–587. [CrossRef]

7. Hock, C.; Villringer, K.; Heekeren, H.; Schuh-Hofer, S.; Hofmann, M.; Minoshima, S.; Schwaiger, M.; Dirnagl, U.; Villringer, A. Decrease in Parietal Cerebral Hemoglobin Oxygenation during Performance of a Verbal Fluency Task in Patients with Alzheimer's Disease Monitored by Means of near-Infrared Spectroscopy (NIRS)—Correlation with Simultaneous RCBF-PET Measurements. *Brain Res.* **1997**, *755*, 293–303. [CrossRef]

8. Fallgatter, A.J.; Roesler, M.; Sitzmann, L.; Heidrich, A.; Mueller, T.J.; Strik, W.K. Loss of Functional Hemispheric Asymmetry in Alzheimer's Dementia Assessed with near-Infrared Spectroscopy. *Cogn. Brain Res.* **1997**, *6*, 67–72. [CrossRef]

9. Arai, H.; Takano, M.; Miyakawa, K.; Ota, T.; Takahashi, T.; Asaka, H.; Kawaguchi, T. A Quantitative Near-Infrared Spectroscopy Study: A Decrease in Cerebral Hemoglobin Oxygenation in Alzheimer's Disease and Mild Cognitive Impairment. *Brain Cogn.* **2006**, *61*, 189–194. [CrossRef] [PubMed]

10. Yang, D.; Huang, R.; Yoo, S.-H.; Shin, M.-J.; Yoon, J.A.; Shin, Y.-I.; Hong, K.-S. Detection of Mild Cognitive Impairment Using Convolutional Neural Network: Temporal-Feature Maps of Functional near-Infrared Spectroscopy. *Front. Aging Neurosci.* **2020**, *12*, 141. [CrossRef]

11. Nguyen, T.; Kim, M.; Gwak, J.; Lee, J.J.; Choi, K.Y.; Lee, K.H.; Kim, J.G. Investigation of Brain Functional Connectivity in Patients with Mild Cognitive Impairment: A Functional Near-infrared Spectroscopy (FNIRS) Study. *J. Biophotonics* **2019**, *12*, e201800298. [CrossRef]

12. Ho, T.K.K.; Kim, M.; Jeon, Y.; Kim, B.C.; Kim, J.G.; Lee, K.H.; Song, J.-I.; Gwak, J. Deep Learning-Based Multilevel Classification of Alzheimer's Disease Using Non-Invasive Functional near-Infrared Spectroscopy. *Front. Aging Neurosci.* **2022**, *14*, 810125. [CrossRef] [PubMed]

13. Niu, H.; Li, X.; Chen, Y.; Ma, C.; Zhang, J.; Zhang, Z. Reduced Frontal Activation during a Working Memory Task in Mild Cognitive Impairment: A Non-invasive Near-infrared Spectroscopy Study. *CNS Neurosci. Ther.* **2013**, *19*, 125–131. [CrossRef]

14. Lee, M.J.; Park, B.-Y.; Cho, S.; Park, H.; Chung, C.-S. Cerebrovascular Reactivity as a Determinant of Deep White Matter Hyperintensities in Migraine. *Neurology* **2019**, *92*, e342–e350. [CrossRef] [PubMed]

15. Lynch, C.E.; Eisenbaum, M.; Algamal, M.; Balbi, M.; Ferguson, S.; Mouzon, B.; Saltiel, N.; Ojo, J.; Diaz-Arrastia, R.; Mullan, M.; et al. Impairment of Cerebrovascular Reactivity in Response to Hypercapnic Challenge in a Mouse Model of Repetitive Mild Traumatic Brain Injury. *J. Cereb. Blood Flow Metab.* **2021**, *41*, 1362–1378. [CrossRef]

16. Chen, S.-F.; Pan, H.-Y.; Huang, C.-R.; Huang, J.-B.; Tan, T.-Y.; Chen, N.-C.; Hsu, C.-Y.; Chuang, Y.-C. Autonomic Dysfunction Contributes to Impairment of Cerebral Autoregulation in Patients with Epilepsy. *J. Pers. Med.* **2021**, *11*, 313. [CrossRef]

17. Blair, G.W.; Doubal, F.N.; Thrippleton, M.J.; Marshall, I.; Wardlaw, J.M. Magnetic Resonance Imaging for Assessment of Cerebrovascular Reactivity in Cerebral Small Vessel Disease: A Systematic Review. *J. Cereb. Blood Flow Metab.* **2016**, *36*, 833–841. [CrossRef]

18. Lin, A.J.; Koike, M.A.; Green, K.N.; Kim, J.G.; Mazhar, A.; Rice, T.B.; LaFerla, F.M.; Tromberg, B.J. Spatial Frequency Domain Imaging of Intrinsic Optical Property Contrast in a Mouse Model of Alzheimer's Disease. *Ann. Biomed. Eng.* **2011**, *39*, 1349–1357. [CrossRef] [PubMed]

19. Lin, A.J.; Liu, G.; Castello, N.A.; Yeh, J.J.; Rahimian, R.; Lee, G.; Tsay, V.; Durkin, A.J.; Choi, B.; LaFerla, F.M.; et al. Optical Imaging in an Alzheimer's Mouse Model Reveals Amyloid-β-Dependent Vascular Impairment. *Neurophotonics* **2014**, *1*, 011005. [CrossRef]

20. Kimura, R.; Ohno, M. Impairments in Remote Memory Stabilization Precede Hippocampal Synaptic and Cognitive Failures in 5XFAD Alzheimer Mouse Model. *Neurobiol. Dis.* **2009**, *33*, 229–235. [CrossRef]

21. Giannoni, P.; Arango-Lievano, M.; Neves, I.D.; Rousset, M.-C.; Baranger, K.; Rivera, S.; Jeanneteau, F.; Claeysen, S.; Marchi, N. Cerebrovascular Pathology during the Progression of Experimental Alzheimer's Disease. *Neurobiol. Dis.* **2016**, *88*, 107–117. [CrossRef] [PubMed]

22. Szu, J.I.; Obenaus, A. Cerebrovascular Phenotypes in Mouse Models of Alzheimer's Disease. *J. Cereb. Blood Flow Metab.* **2021**, *41*, 1821–1841. [CrossRef] [PubMed]

23. Tataryn, N.M.; Singh, V.; Dyke, J.P.; Berk-Rauch, H.E.; Clausen, D.M.; Aronowitz, E.; Norris, E.H.; Strickland, S.; Ahn, H.J. Vascular Endothelial Growth Factor Associated Dissimilar Cerebrovascular Phenotypes in Two Different Mouse Models of Alzheimer's Disease. *Neurobiol. Aging* **2021**, *107*, 96–108. [CrossRef]

24. Seong, M.; Mai, P.M.; Lee, K.; Kim, J.G. Simultaneous Blood Flow and Oxygenation Measurements Using an Off-the-Shelf Spectrometer. *Chin. Opt. Lett.* **2018**, *16*, 071701. [CrossRef]

25. Seong, M.; Phillips, Z.; Mai, P.M.; Yeo, C.; Song, C.; Lee, K.; Kim, J.G. Simultaneous Blood Flow and Blood Oxygenation Measurements Using a Combination of Diffuse Speckle Contrast Analysis and Near-Infrared Spectroscopy. *J. Biomed. Opt.* **2016**, *21*, 027001. [CrossRef]

26. Kim, J.G.; Lee, J.; Mahon, S.B.; Mukai, D.; Patterson, S.E.; Boss, G.R.; Tromberg, B.J.; Brenner, M. Noninvasive Monitoring of Treatment Response in a Rabbit Cyanide Toxicity Model Reveals Differences in Brain and Muscle Metabolism. *J. Biomed. Opt.* **2012**, *17*, 105005. [CrossRef]

27. Prahl, S.; Jacques, S.; OMLC. Available online: https://omlc.org/ (accessed on 29 September 2020).

28. Lee, S.; Kim, J.G. Breast Tumor Hemodynamic Response during a Breath-Hold as a Biomarker to Predict Chemotherapeutic Efficacy: Preclinical Study. *J. Biomed. Opt.* **2018**, *23*, 048001. [CrossRef]

29. Choi, S.S.; Mandelis, A.; Guo, X.; Lashkari, B.; Kellnberger, S.; Ntziachristos, V. Wavelength-Modulated Differential Photoacoustic Spectroscopy (WM-DPAS) for Noninvasive Early Cancer Detection and Tissue Hypoxia Monitoring. *J. Biophotonics* **2016**, *9*, 388–395. [CrossRef]

30. Dehaes, M.; Gagnon, L.; Lesage, F.; Pélégrini-Issac, M.; Vignaud, A.; Valabrègue, R.; Grebe, R.; Wallois, F.; Benali, H. Quantitative Investigation of the Effect of the Extra-Cerebral Vasculature in Diffuse Optical Imaging: A Simulation Study. *Biomed. Opt. Express* **2011**, *2*, 680. [CrossRef]

31. Fang, Q.; Boas, D.A. Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units. *Opt. Express* **2009**, *17*, 20178. [CrossRef]

32. Lee, S.Y.; Zheng, C.; Brothers, R.; Buckley, E.M. Small Separation Frequency-Domain near-Infrared Spectroscopy for the Recovery of Tissue Optical Properties at Millimeter Depths. *Biomed. Opt. Express* **2019**, *10*, 5362. [CrossRef] [PubMed]

33. Choi, J.; Wolf, M.; Toronov, V.; Wolf, U.; Polzonetti, C.; Hueber, D.; Safonova, L.P.; Gupta, R.; Michalos, A.; Mantulin, W.; et al. Noninvasive Determination of the Optical Properties of Adult Brain: Near-Infrared Spectroscopy Approach. *J. Biomed. Opt.* **2004**, *9*, 221–229. [CrossRef]

34. Brigadoi, S.; Cooper, R.J. How Short Is Short? Optimum Source–Detector Distance for Short-Separation Channels in Functional near-Infrared Spectroscopy. *Neurophotonics* **2015**, *2*, 025005. [CrossRef]

35. Fu, X.; Richards, J.E. Age-Related Changes in Diffuse Optical Tomography Sensitivity Profiles in Infancy. *PLoS ONE* **2021**, *16*, e0252036. [CrossRef]

36. Hong, G.; Diao, S.; Chang, J.; Antaris, A.L.; Chen, C.; Zhang, B.; Zhao, S.; Atochin, D.N.; Huang, P.L.; Andreasson, K.I.; et al. Through-Skull Fluorescence Imaging of the Brain in a New near-Infrared Window. *Nat. Photonics* **2014**, *8*, 723–730. [CrossRef] [PubMed]

37. Firbank, M.; Hiraoka, M.; Essenpreis, M.; Delpy, D.T. Measurement of the Optical Properties of the Skull in the Wavelength Range 650-950 Nm. *Phys. Med. Biol.* **1993**, *38*, 503–510. [CrossRef]

38. Genina, E.A.; Bashkatov, A.N.; Tuchin, V.V. Optical Clearing of Cranial Bone. *Adv. Opt. Technol.* **2008**, *2008*, 267867. [CrossRef] [PubMed]

39. Sun, J.; Lee, S.J.; Wu, L.; Sarntinoranont, M.; Xie, H. Refractive Index Measurement of Acute Rat Brain Tissue Slices Using Optical Coherence Tomography. *Opt. Express* **2012**, *20*, 1084. [CrossRef]

40. Jacques, S.L. Optical Properties of Biological Tissues: A Review. *Phys. Med. Biol.* **2013**, *58*, R37–R61. [CrossRef]

41. Gunther, J.E.; Lim, E.A.; Kim, H.K.; Flexman, M.; Altoé, M.; Campbell, J.A.; Hibshoosh, H.; Crew, K.D.; Kalinsky, K.; Hershman, D.L.; et al. Dynamic Diffuse Optical Tomography for Monitoring Neoadjuvant Chemotherapy in Patients with Breast Cancer. *Radiology* **2018**, *287*, 778–786. [CrossRef]

42. Karahaliou, A.; Vassiou, K.; Arikidis, N.S.; Skiadopoulos, S.; Kanavou, T.; Costaridou, L. Assessing Heterogeneity of Lesion Enhancement Kinetics in Dynamic Contrast-Enhanced MRI for Breast Cancer Diagnosis. *Br. J. Radiol.* **2010**, *83*, 296–309. [CrossRef]

43. Kim, J.G.; Liu, H. Investigation of Bi-Phasic Tumor Oxygen Dynamics Induced by Hyperoxic Gas Intervention: A Numerical Study. *Opt. Express* **2005**, *13*, 4465. [CrossRef]

44. Thadewald, T.; Büning, H. Jarque–Bera Test and Its Competitors for Testing Normality—A Power Comparison. *J. Appl. Stat.* **2007**, *34*, 87–105. [CrossRef]

45. Cressie, N.A.C.; Whitford, H.J. How to Use the Two Sample T-Test. *Biom. J.* **1986**, *28*, 131–148. [CrossRef]

46. Sawilowsky, S.S. Misconceptions Leading to Choosing the t Test Over the Wilcoxon Mann-Whitney Test for Shift in Location Parameter. *J. Mod. Appl. Stat. Methods* **2005**, *4*, 598–600. [CrossRef]

47. Ali, M. PyCaret: An Open Source, Low-Code Machine Learning Library in Python. Available online: https://www.pycaret.org (accessed on 1 April 2022).

48. DeMaris, A. A Tutorial in Logistic Regression. *J. Marriage Fam.* **1995**, *57*, 956–968. [CrossRef]

49. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *arXiv* **2012**, arXiv:1201.0490v4. [CrossRef]

50. Tharwat, A.; Gaber, T.; Ibrahim, A.; Hassanien, A.E. Linear Discriminant Analysis: A Detailed Tutorial. *AI Commun.* **2017**, *30*, 169–190. [CrossRef]

51. Tharwat, A. Linear vs. Quadratic Discriminant Analysis Classifier: A Tutorial. *Int. J. Appl. Pattern Recognit.* **2016**, *3*, 145. [CrossRef]

52. Noble, W.S. What Is a Support Vector Machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef]

53. Safavian, S.R.; Landgrebe, D. A Survey of Decision Tree Classifier Methodology. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 660–674. [CrossRef]

54. Bentéjac, C.; Csörgő, A.; Martínez-Muñoz, G. A Comparative Analysis of Gradient Boosting Algorithms. *Artif. Intell. Rev.* **2021**, *54*, 1937–1967. [CrossRef]

55. Pal, M. Random Forest Classifier for Remote Sensing Classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]

56. Schapire, R.E. Explaining AdaBoost. In *Empirical Inference*; Schölkopf, B., Luo, Z., Vovk, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 37–52. ISBN 978-3-642-41135-9.

57. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]

58. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-Sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]

59. Raschka, S. An Overview of General Performance Metrics of Binary Classifer Systems. *arXiv* **2014**, arXiv:1410.5330. [CrossRef]

60. Gaidica, M. Mouse Brain Atlas. Available online: http://labs.gaidi.ca/mouse-brain-atlas/ (accessed on 11 November 2022).

61. Paxinos, G.; Franklin, K.B. *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates*; Academic Press: Cambridge, MA, USA, 2019.

62. Abdi, H.; Williams, L.J. Principal Component Analysis. *Wiley Interdisc. Rev. Comput. Stat.* **2010**, *2*, 433–459. [CrossRef]

63. Ideguchi, H.; Ichiyasu, H.; Fukushima, K.; Okabayashi, H.; Akaike, K.; Hamada, S.; Nakamura, K.; Hirosako, S.; Kohrogi, H.; Sakagami, T.; et al. Validation of a Breath-Holding Test as a Screening Test for Exercise-Induced Hypoxemia in Chronic Respiratory Diseases. *Chron. Respir. Dis.* **2021**, *18*, 14799731211012965. [CrossRef]

64. Stefani, A.; Sancesario, G.; Pierantozzi, M.; Leone, G.; Galati, S.; Hainsworth, A.H.; Diomedi, M. CSF Biomarkers, Impairment of Cerebral Hemodynamics and Degree of Cognitive Decline in Alzheimer's and Mixed Dementia. *J. Neurol. Sci.* **2009**, *283*, 109–115. [CrossRef]

65. Diomedi, M.; Rocco, A.; Bonomi, C.G.; Mascolo, A.P.; De Lucia, V.; Marrama, F.; Sallustio, F.; Koch, G.; Martorana, A. Haemodynamic Impairment along the Alzheimer's Disease Continuum. *Eur. J. Neurol.* **2021**, *28*, 2168–2173. [CrossRef]

66. Brenner, M.; Kim, J.G.; Mahon, S.B.; Lee, J.; Kreuter, K.A.; Blackledge, W.; Mukai, D.; Patterson, S.; Mohammad, O.; Sharma, V.S.; et al. Intramuscular Cobinamide Sulfite in a Rabbit Model of Sublethal Cyanide Toxicity. *Ann. Emerg. Med.* **2010**, *55*, 352–363. [CrossRef] [PubMed]

67. Phillips, Z.; Kim, J.B.; Paik, S.-H.; Kang, S.-Y.; Jeon, N.-J.; Kim, B.-M.; Kim, B.-J. Regional Analysis of Cerebral Hemodynamic Changes during the Head-up Tilt Test in Parkinson's Disease Patients with Orthostatic Intolerance. *Neurophotonics* **2020**, *7*, 045006. [CrossRef] [PubMed]

68. Karamzadeh, N.; Amyot, F.; Kenney, K.; Anderson, A.; Chowdhry, F.; Dashtestani, H.; Wassermann, E.M.; Chernomordik, V.; Boccara, C.; Wegman, E.; et al. A Machine Learning Approach to Identify Functional Biomarkers in Human Prefrontal Cortex for Individuals with Traumatic Brain Injury Using Functional Near-infrared Spectroscopy. *Brain Behav.* **2016**, *6*, e00541. [CrossRef] [PubMed]

69. Yang, D.; Hong, K.-S. Quantitative Assessment of Resting-State for Mild Cognitive Impairment Detection: A Functional Near-Infrared Spectroscopy and Deep Learning Approach. *J. Alzheimers Dis.* **2021**, *80*, 647–663. [CrossRef] [PubMed]

70. Kisler, K.; Nelson, A.R.; Montagne, A.; Zlokovic, B.V. Cerebral Blood Flow Regulation and Neurovascular Dysfunction in Alzheimer Disease. *Nat. Rev. Neurosci.* **2017**, *18*, 419–434. [CrossRef]

71. Singh, C.S.B.; Choi, K.B.; Munro, L.; Wang, H.Y.; Pfeifer, C.G.; Jefferies, W.A. Reversing Pathology in a Preclinical Model of Alzheimer's Disease by Hacking Cerebrovascular Neoangiogenesis with Advanced Cancer Therapeutics. *EBioMedicine* **2021**, *71*, 103503. [CrossRef] [PubMed]

72. Dorr, A.; Sahota, B.; Chinta, L.V.; Brown, M.E.; Lai, A.Y.; Ma, K.; Hawkes, C.A.; McLaurin, J.; Stefanovic, B. Amyloid-β-Dependent Compromise of Microvascular Structure and Function in a Model of Alzheimer's Disease. *Brain* **2012**, *135*, 3039–3050. [CrossRef] [PubMed]

73. Saager, R.B.; Telleri, N.L.; Berger, A.J. Two-Detector Corrected Near Infrared Spectroscopy (C-NIRS) Detects Hemodynamic Activation Responses More Robustly than Single-Detector NIRS. *NeuroImage* **2011**, *55*, 1679–1685. [CrossRef]

74. Lee, J.; Kim, J.G.; Mahon, S.B.; Mukai, D.; Yoon, D.; Boss, G.R.; Patterson, S.E.; Rockwood, G.; Isom, G.; Brenner, M. Noninvasive Optical Cytochrome c Oxidase Redox State Measurements Using Diffuse Optical Spectroscopy. *J. Biomed. Opt.* **2014**, *19*, 055001. [CrossRef]

75. Maurer, I. A Selective Defect of Cytochrome c Oxidase Is Present in Brain of Alzheimer Disease Patients. *Neurobiol. Aging* **2000**, *21*, 455–462. [CrossRef]

76. Cardoso, S.M.; Proença, M.T.; Santos, S.; Santana, I.; Oliveira, C.R. Cytochrome c Oxidase Is Decreased in Alzheimer's Disease Platelets. *Neurobiol. Aging* **2004**, *25*, 105–110. [CrossRef]