

Article

An Isolated CNN Architecture for Classification of Finger-Tapping Tasks Using Initial Dip Images: A Functional Near-Infrared Spectroscopy Study

Muhammad Umair Ali ¹, Amad Zafar ^{1,*}, Karam Dad Kallu ², M. Atif Yaqub ³, Haris Masood ⁴, Keum-Shik Hong ^{5,6} and Muhammad Raheel Bhutta ^{7,*}

- ¹ Department of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Republic of Korea; umair@sejong.ac.kr
- ² Department of Robotics and Intelligent Machine Engineering (RIME), School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), H-12, Islamabad 44000, Pakistan; karamdad.kallu@smme.nust.edu.pk
- ³ ICFO-Institut de Ciències Fotòniques the Barcelona Institute of Science and Technology, 08860 Castelldefels, Spain; atif.yaqub@icfo.eu
- ⁴ Electrical Engineering Department, Wah Engineering College, University of Wah, Wah Cantt 47040, Pakistan; haris.masood@wecuw.edu.pk
- ⁵ School of Mechanical Engineering, Pusan National University, 2 Busandaehak-ro, Geumjeong-gu, Busan 46241, Republic of Korea; kshong@pusan.ac.kr
- ⁶ Institute for Future, School of Automation, Qingdao University, Qingdao 266071, China
- ⁷ Department of Electrical and Computer Engineering, University of UTAH Asia Campus, Incheon 21985, Republic of Korea
- * Correspondence: amad@sejong.ac.kr (A.Z.); raheel.bhutta@utah.edu (M.R.B.)

Abstract: This work investigates the classification of finger-tapping task images constructed for the initial dip duration of hemodynamics (HR) associated with the small brain area of the left motor cortex using functional near-infrared spectroscopy (fNIRS). Different layers (i.e., 16-layers, 19-layers, 22-layers, and 25-layers) of isolated convolutional neural network (CNN) designed from scratch are tested to classify the right-hand thumb and little finger-tapping tasks. Functional *t*-maps of finger-tapping tasks (thumb, little) were constructed for various durations (0.5 to 4 s with a uniform interval of 0.5 s) for the initial dip duration using a three gamma functions-based designed HR function. The results show that the 22-layered isolated CNN model yielded the highest classification accuracy of 89.2% with less complexity in classifying the functional *t*-maps of thumb and little fingers associated with the same small brain area using the initial dip. The results further demonstrated that the active brain area of the two tapping tasks from the same small brain area are highly different and well classified using functional *t*-maps of the initial dip (0.5 to 4 s) compared to functional *t*-maps generated for delayed HR (14 s). This study shows that the images constructed for initial dip duration can be helpful in the future for fNIRS-based diagnosis or cortical analysis of abnormal cerebral oxygen exchange in patients.

Keywords: initial dip; neuronal firing; fNIRS; designed HRF; motor cortex; deep learning



Citation: Ali, M.U.; Zafar, A.; Kallu, K.D.; Yaqub, M.A.; Masood, H.; Hong, K.-S.; Bhutta, M.R. An Isolated CNN Architecture for Classification of Finger-Tapping Tasks Using Initial Dip Images: A Functional Near-Infrared Spectroscopy Study. *Bioengineering* **2023**, *10*, 810. <https://doi.org/10.3390/bioengineering10070810>

Academic Editor: Rui Oliveira

Received: 31 May 2023

Revised: 30 June 2023

Accepted: 30 June 2023

Published: 5 July 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/>).

1. Introduction

Recently, functional near-infrared spectroscopy (fNIRS) has been shown to be a viable non-invasive imaging tool for studying brain activity in individuals of various ages [1,2]. The applications of fNIRS range from behavioral and cognitive neuroscience studies [3] to more innovative applications, including the investigation of cortical hemodynamic variations caused by driving [4], climbing [5], aircraft flying [6], decoding of sleep states including drowsiness [7,8], the administration of psychedelics, and interface with computers [9]. Furthermore, fNIRS has shown promising results when used in conjunction

with various imaging modalities such as optical imaging [9], functional magnetic resonance imaging [10], electroencephalography [11–13], magnetoencephalography [14], or others [15–17].

fNIRS measures both absolute and relative concentration variations of oxyhemoglobin ($\text{HbO}/\Delta\text{HbO}$) and deoxyhemoglobin ($\text{HbR}/\Delta\text{HbR}$) using pairs of numerous near-infrared lights in the range of 650 to 1000 nm, which penetrate through the superficial cortical areas [18,19]. The brain's neocortex is activated by stimuli, increasing blood flow, blood volume, $\text{HbO}/\Delta\text{HbO}$, and $\text{HbR}/\Delta\text{HbR}$ in the regional neocortex [20]. The measured fNIRS signals (i.e., ΔHbO and ΔHbR) pass through three durations, including the initial dip (i.e., ΔHbO decreases and ΔHbR increases), the conventional hemodynamic response (HR) (i.e., ΔHbO increases and ΔHbR decreases), and the undershoot [9,21]. As the first increase (or reduction) in ΔHbR (or ΔHbO) results from an increase in metabolism, followed by an increase in cerebral blood flow, the early dip response is thought to be a faster and better spatial localizer of neuronal activity than the delayed conventional HR [22].

In fNIRS, the identification of neuronal activity in a specific brain area requires the calculation of a precise time series from the obtained ΔHbO and ΔHbR data known as canonical hemodynamic response function (cHRF), which may be derived from the t -statistics [23,24]. The cHRF can be modeled using two or three gamma functions. The desired HRF (dHRF) is then generated by convolving the cHRF with the designed experimental protocol [25]. Finally, the active channels are identified by fitting the dHRF on the experimental measured ΔHbO and ΔHbR signals. Using the results of t -statistics, the functional/spatial maps showing brain activity can be drawn [26]. With the advancement of deep learning in recent years, functional/spatial maps can now play an important role in categorizing or identifying brain areas associated with certain activities [27–29].

Deep learning has been effectively applied to fNIRS signals for classification in pre-processing/augmentation, brain–computer interface (BCI) applications, brain response analysis, and diagnosis [30]. In pre-processing/augmentation, Kim et al. [31] employed a U-Net as a convolutional neural network (CNN) architecture for de-noising the time series using the HRF with a minimum-square error of approximately 0.004–0.005. Furthermore, with reduced motion artifacts, their suggested approach correctly estimated the amplitude and shape of HRF. In the case of BCI, feature extraction is a tedious task. However, recently many studies utilized the benefit of automatic feature extraction characteristics of deep learning to eliminate the manual feature extraction process. Tanveer et al. [27] used color map images constructed at various intervals to classify drowsy and alert states using CNN. They achieved 99.3% average accuracy for drowsy/non-drowsy state detection. Another group achieved a high accuracy of 89% for classifying different tasks using deep neural network (DNN) architecture [32]. Similarly, the work proposed in [28] achieved a high classification accuracy of 99.63% for distinguishing between rest, mental arithmetic, and motor imagery tasks using recurrent CNN. Recently, Khalil et al. [33] classified different workload levels using CNN with an accuracy of 94.52%. For brain analysis using deep learning, researchers have employed CNN to investigate the brain's responses due to stimuli, for example, like/dislike, high/low valence, male or female, discrete preference ratings, etc. [34,35].

Deep learning has also demonstrated its effectiveness in fNIRS-based diagnosis in patients [29,36–38]. In study [39], the activation t -maps and channel correlation maps were used to diagnose patients' mild cognitive impairment (MCI). Three mental tasks, including the N-back tasks, Stroop task, and verbal fluency task, and 15 biomarkers were examined. They obtained 90.62% and 85.58% accuracy for activation t -maps and channel correlation maps, respectively. The same group in another study also utilized temporal-feature maps constructed at thirteen-time points to detect MCI in patients with achievable high accuracy [29]. For the N-back tasks, they achieved an accuracy of 98.61% using the slope map of ΔHbO for an interval of 20–60 s. Similarly, the authors in [40] classified the depressed and non-depressed states using pre-trained CNN models (i.e., ResNet 18, AlexNet) and a support vector machine. They utilized manual (i.e., time and frequency

domain features) and channel correlation to train the models. They showed that ResNet 18, AlexNet, and support vector machine models achieved 76%, 90%, and 83% accuracy between depressed and non-depressed states. In another work [41], the authors utilized connectivity maps generated at various time instances from the resting state data with CNN to MCI detection in patients. Recently, Ho et al. [42] employed deep learning to categorize Alzheimer’s disease stages. All the studies mentioned above utilized various temporal window sizes on conventional HR to generate maps or time series for the same or different brain areas, which were input to deep learning architecture. So far, no study has investigated the effect of initial dip duration on functional maps using fNIRS.

In this study, an isolated CNN network with different layers is investigated for the classification of functional images generated for finger-tapping tasks using the initial duration (i.e., initial dip) of the HR; 16-, 19-, 22-, and 25-layer CNN architecture is designed from scratch. Then, the effectiveness of the designed isolated CNN architecture is evaluated using the right-hand thumb and little finger-tapping tasks using fNIRS signals generated from the left motor cortex. The functional maps for the right-hand thumb finger- (RHTF) and little finger- (RHLF) tapping are generated using *t*-statistic for the initial durations (i.e., 0.5 to 4 s with a uniform interval of 0.5 s) and the HR (14 s), which were further classified using the designed isolated CNN models. Finally, this work shows how functional *t*-maps constructed for the initial dip may be used to categorize several tasks performed by the same local brain region.

The paper is divided into the following sections. Section 2 discusses the fNIRS data acquisition and pre-processing, image construction, isolated CNN models, and proposed framework. The results of the proposed framework are presented in Section 3. Finally, Sections 4 and 5 discuss and conclude the findings of this study.

2. Materials and Methods

Figure 1 shows the proposed isolated CNN architecture framework for the classification of RHTF- and RHLF-tapping tasks from the same brain area (i.e., left motor cortex).

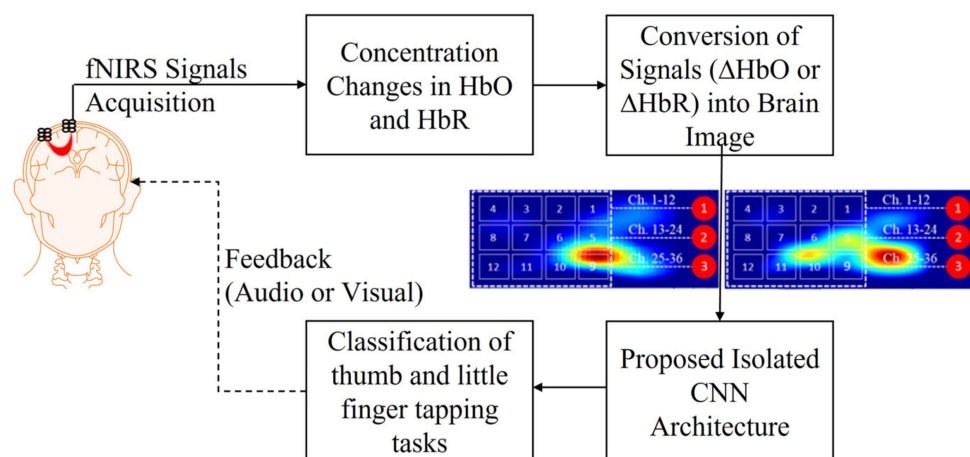


Figure 1. Proposed isolated CNN architecture framework.

2.1. Experimental Data and Pre-Processing

The proposed isolated CNN framework was validated using previously published experimental data of finger tapping (i.e., RHTF and RHLF) from eleven (11) volunteers (age: 28.5 ± 2.5 years, gender: male) [43]. None had a history of neurological or visual disorders; all were in good health with normal or corrected-to-normal vision. The fNIRS data of RHTF- and RHLF-tapping tasks were captured from the left motor cortex at a 9.19 Hz sampling rate using frequency domain ISS Imagent, ISS Inc. (Champaign IL, USA), equipment using near-infrared light of wavelengths 690 nm and 830 nm. The raw intensity data were acquired using ISS Imagent data collection and analysis software (ISS

Boxy). The ISS-Boxy software was then used to convert the intensity data to ΔHbO and ΔHbR using the modified Beer–Lambert law. In a dense-emitter-detector arrangement, thirty-six channels using three detectors and twelve emitters were placed at the C3 area to examine the left motor cortex. The experiment took place in a dark, quiet room. The experimental protocol entails two sessions of tapping activities (i.e., one for RHTF-tapping and the other for RHLF-tapping) with 60 s of pre-rest and 10 s of post-rest. The six trials each last 30 s and make up a session. A 10 s tapping task is followed by a rest duration of 20 s in each trial. The individuals were asked to tap their RHTF and RHLF as quickly as they could without regard for the number of taps throughout the task time. A dark screen was displayed throughout the rest interval. The volunteers were also told to keep their eyes open during the experiment. To reduce physiological noises from respiration, heartbeat, and low-frequency drift, the measured data were pre-processed using a 4th-order band-pass filter having 0.01 and 0.15 Hz cutoff frequencies [44]. Further details on the optode configuration, paradigm protocol, and pre-processing can be found in [43,44].

2.2. Image Construction

The most critical step in analyzing fNIRS data is estimating cortical activity and localizing it. By adjusting the recorded HR to the predetermined dHRF, cortical activity may be determined [45], and its presence may be inferred from the associated channels' t -values. A higher t -value specifies that the measured channel signal (i.e., ΔHbO) and dHRF have a strong correlation. In this paper, dHRF was generated by convolving cHRF made of three gamma functions with the experimental protocol (i.e., 10 s for RHTF or RHLF tasks and 20 s for the rest duration). The use of cHRF with three gamma functions enables the visualization of functional activation of the initial dip duration. In this paper, using MATLAB's robustfit command, the t -values were determined. The activation in the functional maps were decided for t -value $> t_{\text{crt}}$ and p -value < 0.05 or 0 for other t -values. The t_{crt} depends on the degrees of freedom of the data. In this study, the value 1.65 is utilized for t_{crt} (i.e., for 30 s trial, 275 data points ($30\text{ s} \times 9.19\text{ Hz}$), 274 degrees of freedom). In order to show activation in the covered brain area, the resulting t -values were normalized within the 0~1 range and shown on a t -map as a functional activation map. All the images were constructed at 227×227 pixel size. Only ΔHbO data were used for further analysis since several prior research has demonstrated that ΔHbO is a more sensitive and accurate indicator [3,26,45].

In this work, the functional t -maps for RHTF- and RHLF-tapping tasks were drawn for various initial dip durations (i.e., 0.5 s, 1 s, 1.5 s, 2 s, 2.5 s, 3 s, 3.5 s, and 4 s) and at 14 s for comparison with conventional delayed HR using the ΔHbO data. Figure 2 shows the functional t -maps drawn at different durations for RHTF- and RHLF-tapping tasks, respectively. Our hypothesis is that the functional t -maps acquired at the initial dip duration (i.e., 0.5 s, 1 s, 1.5 s, 2 s, 2.5 s, 3 s, 3.5 s, and 4 s) are more spatially precise than that of the functional t -map produced at 14 s and yield better classification accuracy for the tapping tasks acquired from the same brain area.

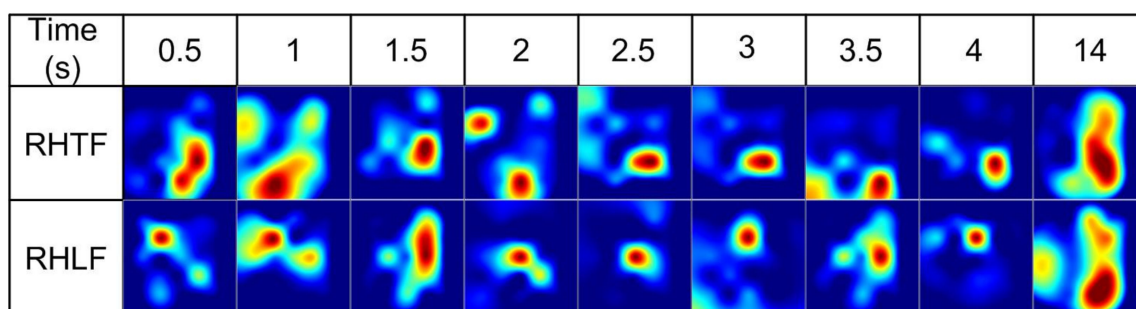
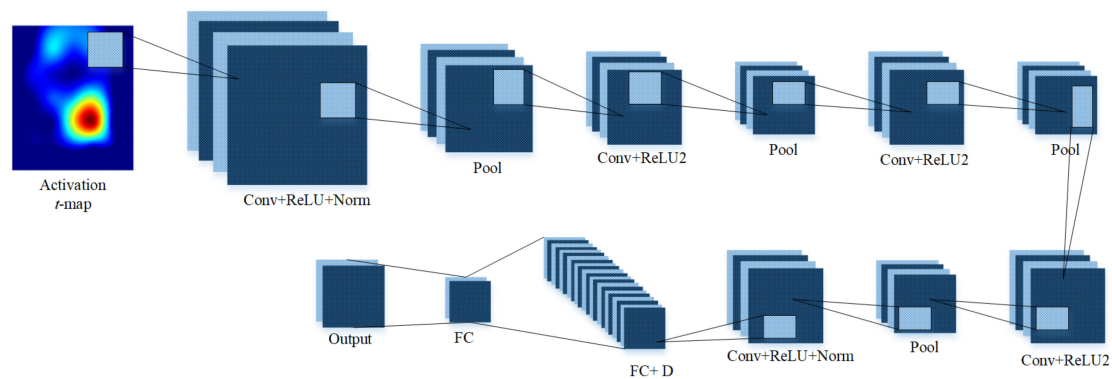


Figure 2. Functional t -maps at various time durations.

2.3. Isolated Convolutional Neural Network

CNN is the most widely used deep learning model in neural networks. The two components of a typical CNN model are feature extraction and classification. The input layer, convolution layer, pooling layer, fully connected layer, and classification layer are the five key layers of a CNN architecture. Through successively arranging trainable layers one after the other, CNN conducts feature extraction and classification. Convolutional and pooling layers are often included in the feature extraction section of a CNN, whereas fully connected and classification layers are typically included in the classification section. CNNs have been widely employed in many different fields whose input data are other than images, such as audio and video [46].

In this work, the four CNN models (16-layers, 19-layers, 22-layers, and 25-layers) are designed to classify the functional *t*-maps of RHTF- and RHTL-tapping tasks from the left motor cortex. The isolated CNN models acquire their input from RHTF and RHTL functional *t*-maps. CNN models with 16-layers, 19-layers, 22-layers, and 25-layers are designed by combining input, convolution, ReLU, normalization, maxpool, dropout, and fully connected layers. Figure 3 illustrates the 22-layered CNN model’s architect with a detailed description of each layer.



Layer Name	Layer Type	Properties
imageinput	Image Input	227×227×3 images with 'zerocenter' normalization
conv_1	Convolution	128 6×6 convolutions with stride [4 4] and padding [0 0 0 0]
relu_1	ReLU	ReLU
crossnorm	Cross Channel Normalization	cross channel normalization with 5 channels per element
maxpool_1	Max	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
conv_2	Convolution	96 6×6 convolutions with stride [1 1] and padding [2 2 2 2]
relu_2	ReLU	ReLU
maxpool_2	Max	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
conv_3	Convolution	96 2×2 convolutions with stride [1 1] and padding [2 2 2 2]
relu_3	ReLU	ReLU
maxpool_3	Max	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
conv_4	Convolution	24 6×6 convolutions with stride [1 1] and padding [2 2 2 2]
relu_4	ReLU	ReLU
maxpool_4	Max	2×2 max pooling with stride [2 2] and padding [0 0 0 0]
conv_5	Convolution	24 6×6 convolutions with stride [1 1] and padding [2 2 2 2]
relu_5	ReLU	ReLU
batchnorm	Batch Normalization	Batch normalization
fc_2	Fully	512 fully connected layer
dropout	Dropout	30% dropout
fc_1	Fully	4 fully connected layer
softmax	Softmax	softmax
classoutput	Classification Output	crossentropyex

Figure 3. Proposed isolated CNN model to classify the functional *t*-maps of tapping tasks.

3. Results

All the processing was conducted using MATLAB™ 2021a. Firstly, the functional *t*-maps of RHTF- and RHLF-tapping tasks generated at various intervals for the initial dip duration (i.e., 0.5–4 s with a uniform interval of 0.5 s) and delayed HR at 14 s are individually fed (i.e., 11 subjects × 6 trials of RHTF/RHLF = 66 RHTF functional *t*-maps and 66 RHLF functional *t*-maps for each interval) to the isolated CNN designed using different layers (i.e., 16-layers, 19-layers, 22-layers, and 25-layers). For all models, 100 epochs, 0.01 initial learn rate, 0.9 momentum, and SGDM solver were used. In order to avoid the over-fitting of data, the data of each activity (i.e., RHTF and RHLF) were divided into a 70:30 ratio (70% data for training and 30% data for validation). Figure 4 shows the accuracy for different-layered isolated CNN architecture at various time durations. Isolated CNN with 22-layers yielded the best accuracy with less complexity. However, the classification accuracy is lower than acceptable (i.e., 70%). In most cases, the classification accuracy is also below the chance level (i.e., 50%).

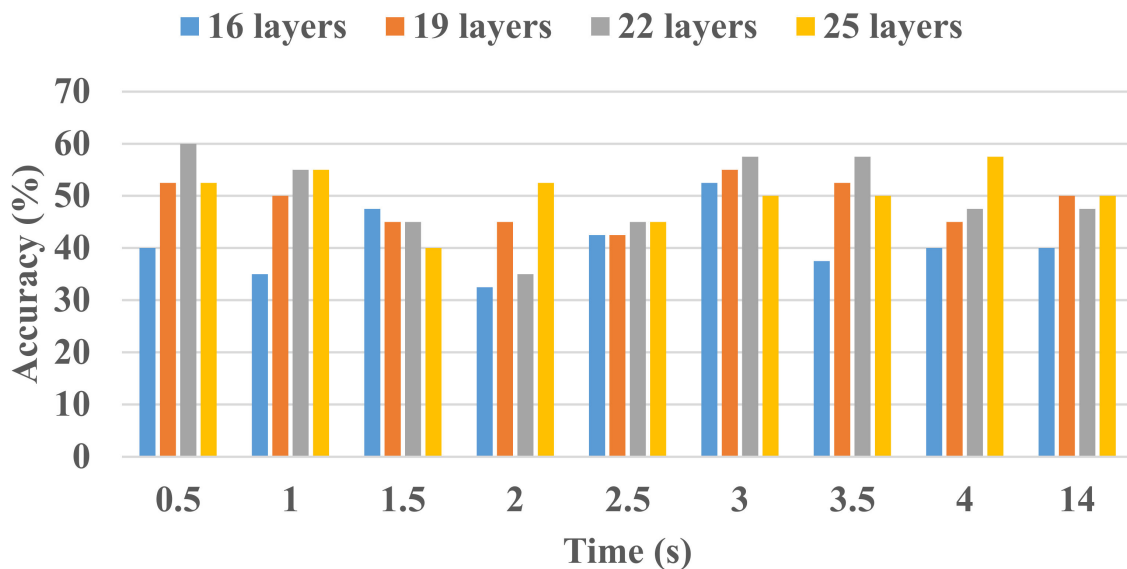


Figure 4. Classification accuracy for different layers' (16-layers, 19-layers, 22-layers, and 25-layers) isolated CNN architecture at various time durations.

To increase the classification accuracy for initial dip duration, functional *t*-maps obtained at various intervals were combined to increase the size of the dataset (i.e., 11 subjects × 6 trials of RHTF/RHLF × 8 intervals = 528 RHTF function *t*-maps and 528 RHLF functional *t*-maps). Out of 528 images of each activity, 370 images were used to train the model, whereas the remaining unseen images were used for the validation of the model. The classification accuracy, loss curves, and validation accuracy of the training stage to distinguish RHTF- and RHLF-tapping tasks of the left motor cortex using the initial dip durations for different layers of isolated CNN are shown in Figures 5–7 respectively. The validation accuracy of 85.8%, 86.1%, 89.2%, and 89.9% were obtained for 16-layers, 19-layers, 22-layers, and 25-layers of isolated CNN models for the two classes (i.e., RHTF- and RHLF-tapping tasks).

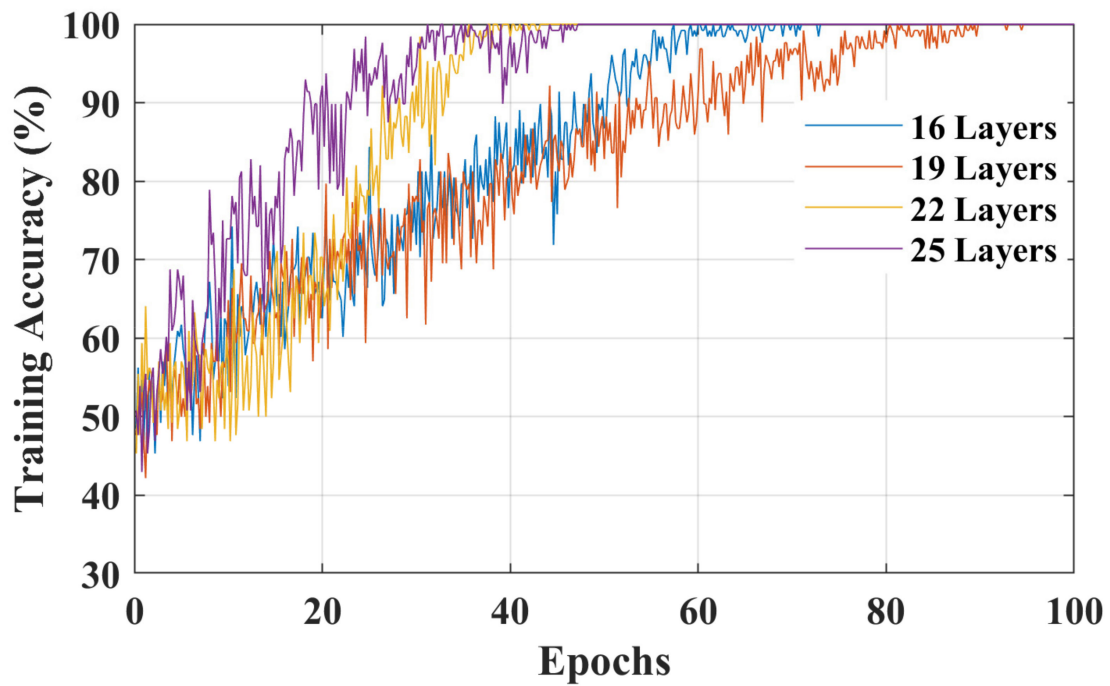


Figure 5. Classification accuracy for 16-layers, 19-layers, 22-layers, and 25-layers of isolated CNN architecture using the functional *t*-maps of initial dip (i.e., 0.5 to 4 s with a uniform interval of 0.5 s).

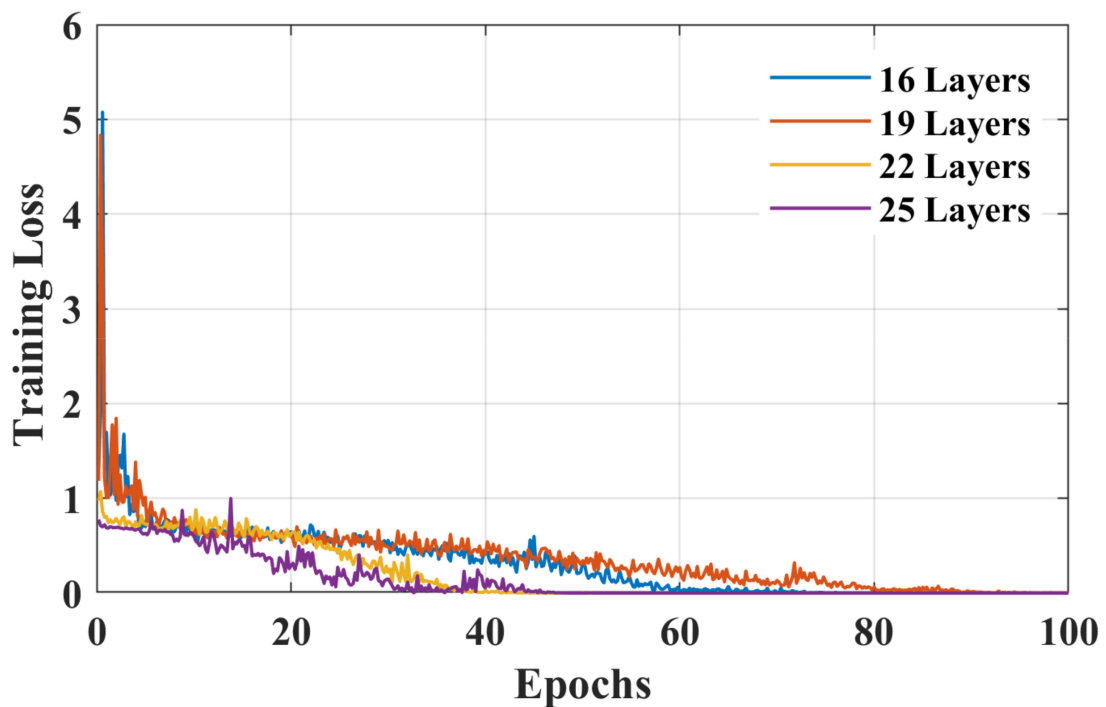


Figure 6. Loss curves for 16-layers, 19-layers, 22-layers, and 25-layers of isolated CNN architecture using the functional *t*-maps of initial dip (i.e., 0.5 to 4 s with a uniform interval of 0.5 s).

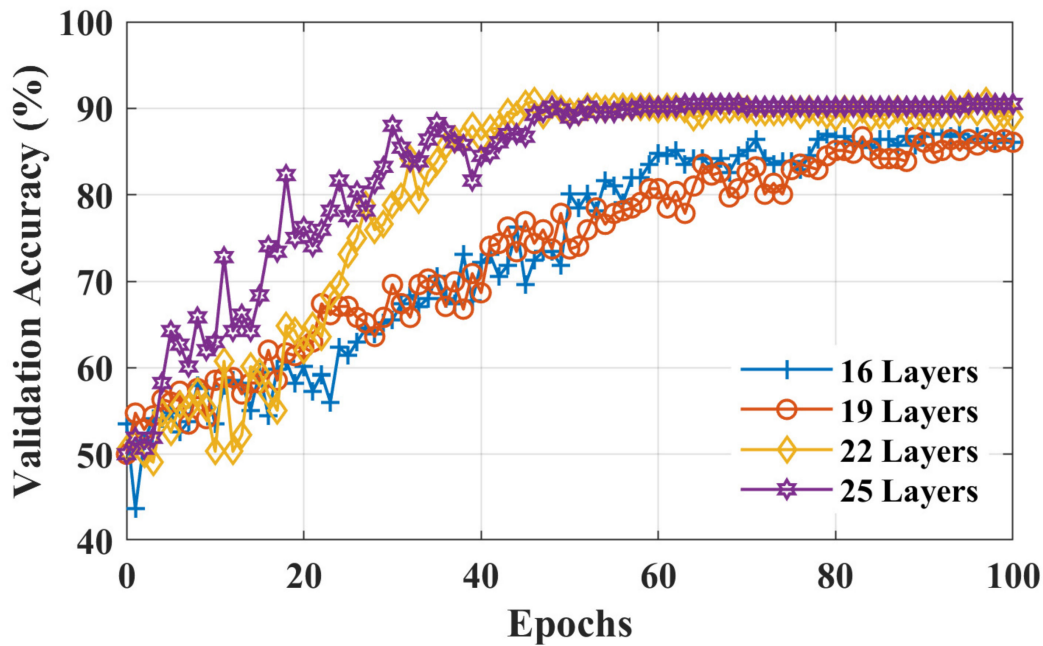


Figure 7. Validation accuracy for 16-layers, 19-layers, 22-layers, and 25-layers of isolated CNN architecture using the functional *t*-maps of initial dip (i.e., 0.5 to 4 s with a uniform interval of 0.5 s).

The 25-layered isolated CNN architecture yielded a classification accuracy of 89.9%, but the complexity of CNN architecture also increased. In comparison to the 25-layered isolated CNN network, it can be noted from Figures 5–7 that the isolated CNN architecture with 22-layers yielded an accuracy of 89.2% for the classification of RHTF- and RHLF-tapping tasks for the initial dip durations from the same small brain area (i.e., left motor cortex) with low complexity and fast convergence. Furthermore, true positive rate (TPR), false negative rate (FNR), positive predictive value (PPV), and false discovery rate (FDR) metrics were calculated using Equation (1) for the selection of network layers, as shown in Table 1.

$$\left. \begin{aligned}
 \text{TPR (\%)} &= \frac{\text{True positive}}{\text{No. of real positive}} \times 100 \\
 \text{FNR (\%)} &= \frac{\text{False negative}}{\text{No. of real positive}} \times 100 \\
 \text{PPV (\%)} &= \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \times 100 \\
 \text{FDR (\%)} &= \frac{\text{False positive}}{\text{True positive} + \text{False positive}} \times 100
 \end{aligned} \right\} \quad (1)$$

Table 1. Performance of various network layers (i.e., 16-layers, 19-layers, 22-layers, and 25-layers) to classify functional *t*-maps of initial dip (i.e., 0.5 to 4 s with a uniform interval of 0.5 s).

Network Layers	Class	Classified as		TPR (%)	FNR (%)	PPV (%)	FDR (%)	Accuracy (%)
		RHLF	RHTF					
16	RHLF	136	22	86.1	13.9	85.5	14.5	85.8
	RHTF	23	135	85.4	14.6	86.0	14.0	
19	RHLF	134	24	84.8	15.2	87.0	13.0	86.1
	RHTF	20	138	87.3	12.7	85.2	14.8	
22	RHLF	140	18	88.6	11.4	89.7	10.3	89.2
	RHTF	16	142	89.9	10.1	88.8	11.3	
25	RHLF	144	14	91.1	8.9	88.9	11.1	89.9
	RHTF	18	140	88.6	11.4	90.9	9.1	

Table 1 shows the confusion matrix of the validation stage for 16-layers, 19-layers, 22-layers, and 25-layers of isolated CNN models. It can be seen that the 22-layered isolated CNN architecture correctly classified 142 unseen images of RHTF and 140 unseen images of RHLF-tapping tasks, respectively. It has a TPR of 88.6% and 89.9% for RHLF and RHTF, respectively, whereas the 25-layered model has less TPR for RHTF than the 22-layered model and has a high TPR for the RHLF class. After carefully analyzing the results, it was noted that the 25-layered model has not significantly increased the model accuracy compared to the 22-layered model, but the complexity of the model was increased. Therefore, the area under the curve (AUC) of the 22-layered isolated CNN architecture (proposed model) is shown in Figure 8.

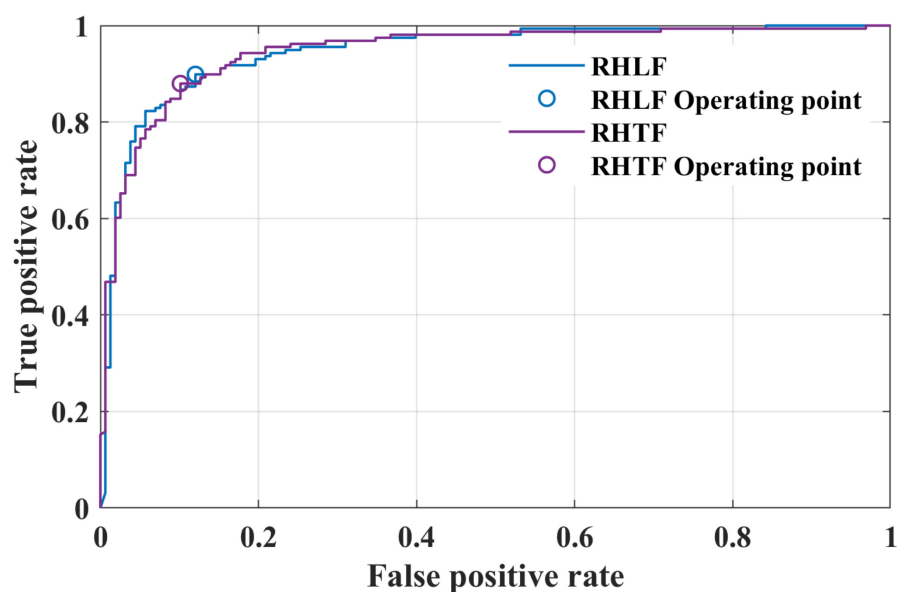


Figure 8. AUC for 22-layered isolated CNN architecture for validation stage using the functional t -maps of initial dip (i.e., 0.5 to 4 s with a uniform interval of 0.5 s).

4. Discussion

The novelties of this work are two-fold: (i) the designing of an isolated CNN network for the fNIRS data and (ii) using the initial dip duration (0.5 to 4 s with a uniform interval of 0.5 s) functional t -maps to classify the finger-tapping tasks from the same small brain area (left motor cortex).

In a previous study by [43], the classification accuracy of $74.9\% \pm 6.4\%$ was obtained for the same RHTF- and RHLF-tapping tasks using the time domain features of the initial dip duration. However, our hypothesis in this study was that the images created using the initial dip duration would give better classification accuracy for the activities obtained from the same small brain area. Furthermore, CNN provides the further advantage of automatic feature extraction by learning the image data and giving better accuracy. In this paper, the functional t -maps were drawn using the three gamma functions. The initial dip was included in the dHRF model with three gamma functions, adding an additional degree of freedom and improving the fNIRS signal's estimate and prediction [26]. The obtained results supported the hypothesis by showing that the classification accuracy is increased to 89.2% for the classification of RHTF- and RHLF-tapping tasks by using the initial dip images (see Figures 5–7). Our findings are in line with the previous literature showing that the initial dip duration reflecting the extraction of oxygen at the task initiation is spatially more specific than the delayed HR [9,22,43,44,47]. In comparison to initial dip images, the classification accuracy decreased using the delayed HR (images drawn at 14 s), which shows that it is difficult to classify the different tasks originating from the small brain area due to an increase in cerebral blood flow [9,22,48,49].

In recent years, deep learning has been successfully applied for pre-processing/augmentation, brain–computer interface (BCI) applications, brain response analysis, and diagnosis using the fNIRS signals [30]. However, all aforementioned studies focused on the conventional delayed HR (different temporal windows) obtained for various activities from the different brain areas [28,29,31,39,41,42]. In comparison, this study is focused on the initial dip duration (0.5 to 4 s with a uniform interval of 0.5 s) and classification of finger-tapping tasks (thumb and little) from the same small brain area. Furthermore, in this study, different layers of isolated CNN architecture (i.e., 16-layers, 19-layers, 22-layers, and 25-layers) were tested for fNIRS signals (thumb- and little finger-tapping) obtained from the left motor cortex. Initially, all networks yielded low accuracy when trained with images obtained at various time durations (i.e., 0.5–4 s with an interval of 0.5 s and 14 s). The reason might be the small size of images (i.e., 66 images for each task) and fewer details for the initial dip duration. Several previous studies suggest that the initial dip duration peaks at 2 s and are complete at 4 s [9,50–52]. Therefore, the size of the dataset was increased by combining all images obtained at various duration from 0.5 to 4 s (i.e., 528 images for each task). The results show that the isolated CNN trained well using the images of different durations, enabling the proposed model to yield an accuracy of 89.2% for unseen images (see Table 1). Furthermore, Figure 8 also shows that both tasks were classified with almost similar classification rates. Our result suggests that the brain response due to different tasks originating in the same small brain area can be well classified using the initial dip images.

Finally, this study demonstrates that the images obtained for the initial dip duration are spatially more specific for the activities originating from the same small brain area. The initial dip duration image depicts the fast oxygen extraction at the start of the task initiation. These results will be helpful in the future for diagnosis or critical analysis of patients with brain injuries. The limitation of this study is that only functional *t*-maps and ΔHbO were used for analysis. Therefore, other markers for generating functional images should be investigated in the future. Furthermore, the analysis using ΔHbR , change in cerebral oxygen exchange (ΔCOE), and change in cerebral blood volume (ΔCBV) should be incorporated to enhance the classification accuracy for more accurate cortical analysis.

5. Conclusions

In this study, initial dip-based functional images and isolated CNNs were used to classify the right-hand finger-tapping tasks (thumb and little) of the same brain area (left motor cortex) using fNIRS signals. Various layers (i.e., 16-layers, 19-layers, 22-layers, and 25-layers) of isolated CNN were designed. All isolated CNN models were trained on the functional *t*-maps of the right hand's thumb- and little finger-tapping tasks generated using three gamma functions-based designed HR; the 22-layered CNN yielded the highest accuracy. The results demonstrated that the two tapping tasks of the same brain area are quite distinguishable, with an accuracy of 89.2% using the initial dip-based functional *t*-maps drawn for 0.5 to 4 s with a uniform interval of 0.5 s. This study shows that the brain response due to different tasks originating in the same small brain area can be well classified using the initial dip images and can be useful in the future for fNIRS-based diagnosis or cortical analysis.

Author Contributions: Conceptualization, M.U.A. and A.Z.; Methodology, M.U.A. and M.R.B.; Software, H.M.; Validation, A.Z.; Formal analysis, K.D.K.; Investigation, M.A.Y.; Resources, K.-S.H.; Writing—original draft, M.U.A. and A.Z.; Writing—review and editing, H.M.; K.-S.H.; and M.R.B.; Supervision, K.-S.H.; Project administration, K.-S.H.; Funding acquisition, M.R.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Research Foundation (NRF) of Korea through the auspices of the Ministry of Science and ICT, Republic of Korea, under Grant NRF-2020R1G1A1012741 (M.R.B.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are included in the article.

Acknowledgments: This research was funded by the National Research Foundation (NRF) of Korea through the auspices of the Ministry of Science and ICT, Republic of Korea, under Grant NRF-2020R1G1A1012741 (M.R.B.)

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

References

1. Karen, T.; Kleiser, S.; Ostojic, D.; Isler, H.; Guglielmini, S.; Bassler, D.; Wolf, M.; Scholkmann, F. Cerebral hemodynamic responses in preterm-born neonates to visual stimulation: Classification according to subgroups and analysis of frontotemporal–occipital functional connectivity. *Neurophotonics* **2019**, *6*, 045005. [[CrossRef](#)] [[PubMed](#)]
2. Llana, T.; Fernandez-Baizan, C.; Mendez-Lopez, M.; Fidalgo, C.; Mendez, M. Functional near-infrared spectroscopy in the neuropsychological assessment of spatial memory: A systematic review. *Acta Psychol.* **2022**, *224*, 103525. [[CrossRef](#)]
3. Bonetti, L.V.; Hassan, S.A.; Lau, S.-T.; Melo, L.T.; Tanaka, T.; Patterson, K.K.; Reid, W.D. Oxyhemoglobin changes in the prefrontal cortex in response to cognitive tasks: A systematic review. *Int. J. Neurosci.* **2019**, *129*, 194–202. [[CrossRef](#)]
4. Yoshino, K.; Oka, N.; Yamamoto, K.; Takahashi, H.; Kato, T. Correlation of prefrontal cortical activation with changing vehicle speeds in actual driving: A vector-based functional near-infrared spectroscopy study. *Front. Hum. Neurosci.* **2013**, *7*, 895. [[CrossRef](#)]
5. Carius, D.; Hörnig, L.; Ragert, P.; Kaminski, E. Characterizing cortical hemodynamic changes during climbing and its relation to climbing expertise. *Neurosci. Lett.* **2020**, *715*, 134604. [[CrossRef](#)] [[PubMed](#)]
6. Gateau, T.; Durantin, G.; Lancelot, F.; Scannella, S.; Dehais, F. Real-time state estimation in a flight simulator using fNIRS. *PLoS ONE* **2015**, *10*, e0121279. [[CrossRef](#)]
7. Delaire, É.; Abdallah, C.; Uji, M.; Cai, Z.; Brooks, M.; Minato, E.; Mozhentiy, E.; Spilkin, A.; Keraudran, H.; Bakian, S. Personalized EEG/fNIRS: A promising tool to study whole-night sleep in healthy and pathological conditions. *Sleep Med.* **2022**, *100*, S21–S22. [[CrossRef](#)]
8. Shoaib, Z.; Akbar, A.; Kim, E.S.; Kamran, M.A.; Kim, J.H.; Jeong, M.Y. Utilizing EEG and fNIRS for the detection of sleep-deprivation-induced fatigue and its inhibition using colored light stimulation. *Sci. Rep.* **2023**, *13*, 6465. [[CrossRef](#)]
9. Hong, K.-S.; Zafar, A. Existence of initial dip for BCI: An illusion or reality. *Front. Neurobot.* **2018**, *12*, 69. [[CrossRef](#)]
10. Scarapicchia, V.; Brown, C.; Mayo, C.; Gawryluk, J.R. Functional magnetic resonance imaging and functional near-infrared spectroscopy: Insights from combined recording studies. *Front. Hum. Neurosci.* **2017**, *11*, 419. [[CrossRef](#)]
11. Chiarelli, A.M.; Zappasodi, F.; Di Pompeo, F.; Merla, A. Simultaneous functional near-infrared spectroscopy and electroencephalography for monitoring of human brain activity and oxygenation: A review. *Neurophotonics* **2017**, *4*, 041411. [[CrossRef](#)]
12. Liu, Z.; Shore, J.; Wang, M.; Yuan, F.; Buss, A.; Zhao, X. A systematic review on hybrid EEG/fNIRS in brain-computer interface. *Biomed. Signal Process. Control* **2021**, *68*, 102595. [[CrossRef](#)]
13. Gao, F.; Hua, L.; He, Y.; Xu, J.; Li, D.; Zhang, J.; Yuan, Z. Word Structure Tunes Electrophysiological and Hemodynamic Responses in the Frontal Cortex. *Bioengineering* **2023**, *10*, 288. [[CrossRef](#)] [[PubMed](#)]
14. Huppert, T.; Barker, J.; Schmidt, B.; Walls, S.; Ghuman, A. Comparison of group-level, source localized activity for simultaneous functional near-infrared spectroscopy-magnetoencephalography and simultaneous fNIRS-fMRI during parametric median nerve stimulation. *Neurophotonics* **2017**, *4*, 015001. [[CrossRef](#)] [[PubMed](#)]
15. Chen, W.-L.; Wagner, J.; Heugel, N.; Sugar, J.; Lee, Y.-W.; Conant, L.; Malloy, M.; Heffernan, J.; Quirk, B.; Zinos, A. Functional near-infrared spectroscopy and its clinical application in the field of neuroscience: Advances and future directions. *Front. Neurosci.* **2020**, *14*, 724. [[CrossRef](#)]
16. Ali, M.U.; Kim, K.S.; Kallu, K.D.; Zafar, A.; Lee, S.W. OptEF-BCI: An Optimization-Based Hybrid EEG and fNIRS-Brain Computer Interface. *Bioengineering* **2023**, *10*, 608. [[CrossRef](#)]
17. Shoaib, Z.; Chang, W.K.; Lee, J.; Lee, S.H.; Phillips, V.Z.; Lee, S.H.; Paik, N.-J.; Hwang, H.-J.; Kim, W.-S. Investigation of neuromodulatory effect of anodal cerebellar transcranial direct current stimulation on the primary motor cortex using functional near-infrared spectroscopy. *Cerebellum* **2023**, 1–11. [[CrossRef](#)] [[PubMed](#)]
18. Boas, D.A.; Elwell, C.E.; Ferrari, M.; Taga, G. Twenty years of functional near-infrared spectroscopy: Introduction for the special issue. *Neuroimage* **2014**, *85*, 1–5. [[CrossRef](#)]
19. Scholkmann, F.; Kleiser, S.; Metz, A.J.; Zimmermann, R.; Pavia, J.M.; Wolf, U.; Wolf, M. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* **2014**, *85*, 6–27. [[CrossRef](#)]
20. Reganova, E.; Solovyeva, K.; Buyanov, D.; Gerasimenko, A.Y.; Repin, D. Effects of Intermittent Hypoxia and Electrical Muscle Stimulation on Cognitive and Physiological Metrics. *Bioengineering* **2023**, *10*, 536. [[CrossRef](#)]
21. Ernst, T.; Hennig, J. Observation of a fast response in functional MR. *Magn. Reson. Med.* **1994**, *32*, 146–149. [[CrossRef](#)]
22. Duong, T.Q.; Kim, D.S.; Uğurbil, K.; Kim, S.G. Spatiotemporal dynamics of the BOLD fMRI signals: Toward mapping submillimeter cortical columns using the early negative response. *Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med.* **2000**, *44*, 231–242. [[CrossRef](#)]

23. Abdelnour, A.F.; Huppert, T. Real-time imaging of human brain function by near-infrared spectroscopy using an adaptive general linear model. *Neuroimage* **2009**, *46*, 133–143. [[CrossRef](#)] [[PubMed](#)]
24. Huppert, T.J.; Diamond, S.G.; Franceschini, M.A.; Boas, D.A. HomER: A review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl. Opt.* **2009**, *48*, D280–D298. [[CrossRef](#)]
25. Lindquist, M.A.; Loh, J.M.; Atlas, L.Y.; Wager, T.D. Modeling the hemodynamic response function in fMRI: Efficiency, bias and mis-modeling. *Neuroimage* **2009**, *45*, S187–S198. [[CrossRef](#)]
26. Ye, J.C.; Tak, S.; Jang, K.E.; Jung, J.; Jang, J. NIRS-SPM: Statistical parametric mapping for near-infrared spectroscopy. *Neuroimage* **2009**, *44*, 428–447. [[CrossRef](#)]
27. Tanveer, M.A.; Khan, M.J.; Qureshi, M.J.; Naseer, N.; Hong, K.-S. Enhanced drowsiness detection using deep learning: An fNIRS study. *IEEE Access* **2019**, *7*, 137920–137929. [[CrossRef](#)]
28. Ghonchi, H.; Fateh, M.; Abolghasemi, V.; Ferdowsi, S.; Rezvani, M. Deep recurrent-convolutional neural network for classification of simultaneous EEG-fNIRS signals. *IET Signal Process.* **2020**, *14*, 142–153. [[CrossRef](#)]
29. 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](#)] [[PubMed](#)]
30. Eastmond, C.; Subedi, A.; De, S.; Intes, X. Deep Learning in fNIRS: A review. *arXiv* **2022**, arXiv:2201.13371. [[CrossRef](#)] [[PubMed](#)]
31. Kim, M.; Lee, S.; Dan, I.; Tak, S. A deep convolutional neural network for estimating hemodynamic response function with reduction of motion artifacts in fNIRS. *J. Neural Eng.* **2022**, *19*, 016017. [[CrossRef](#)] [[PubMed](#)]
32. Saadati, M.; Nelson, J.; Ayaz, H. Multimodal fNIRS-EEG classification using deep learning algorithms for brain-computer interfaces purposes. In Proceedings of the International Conference on Applied Human Factors and Ergonomics, Washington, DC, USA, 24–28 July 2019; pp. 209–220.
33. Khalil, K.; Asgher, U.; Ayaz, Y. Novel fNIRS study on homogeneous symmetric feature-based transfer learning for brain-computer interface. *Sci. Rep.* **2022**, *12*, 1–12. [[CrossRef](#)] [[PubMed](#)]
34. Ramirez, M.; Kaheh, S.; Khalil, M.A.; George, K. Application of Convolutional Neural Network for Classification of Consumer Preference from Hybrid EEG and FNIRS Signals. In Proceedings of the 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 26–29 January 2022; pp. 1024–1028.
35. Bandara, D.; Hirshfield, L.; Velipasalar, S. Classification of affect using deep learning on brain blood flow data. *J. Near Infrared Spectrosc.* **2019**, *27*, 206–219. [[CrossRef](#)]
36. Xu, L.; Geng, X.; He, X.; Li, J.; Yu, J. Prediction in autism by deep learning short-time spontaneous hemodynamic fluctuations. *Front. Neurosci.* **2019**, *13*, 1120. [[CrossRef](#)]
37. Sirpal, P.; Kassab, A.; Pouliot, P.; Nguyen, D.K.; Lesage, F. fNIRS improves seizure detection in multimodal EEG-fNIRS recordings. *J. Biomed. Opt.* **2019**, *24*, 051408. [[CrossRef](#)]
38. Khan, M.A.; Ghafoor, U.; Yoo, H.-R.; Hong, K.-S. Evidence of Neuroplasticity Due to Acupuncture: An fNIRS Study. *Preprint* **2020**. [[CrossRef](#)]
39. Yang, D.; Hong, K.-S.; Yoo, S.-H.; Kim, C.-S. Evaluation of neural degeneration biomarkers in the prefrontal cortex for early identification of patients with mild cognitive impairment: An fNIRS study. *Front. Hum. Neurosci.* **2019**, *13*, 317. [[CrossRef](#)]
40. Wang, R.; Hao, Y.; Yu, Q.; Chen, M.; Humar, I.; Fortino, G. Depression analysis and recognition based on functional near-infrared spectroscopy. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4289–4299. [[CrossRef](#)]
41. 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. Alzheimer's Dis.* **2021**, *80*, 647–663. [[CrossRef](#)]
42. 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*. [[CrossRef](#)]
43. Zafar, A.; Hong, K.-S. Neuronal activation detection using vector phase analysis with dual threshold circles: A functional near-infrared spectroscopy study. *Int. J. Neural Syst.* **2018**, *28*, 1850031. [[CrossRef](#)]
44. Zafar, A.; Hong, K.-S. Reduction of onset delay in functional near-infrared spectroscopy: Prediction of HbO/HbR signals. *Front. Neuroinformatics* **2020**, *14*, 10. [[CrossRef](#)]
45. Tak, S.; Ye, J.C. Statistical analysis of fNIRS data: A comprehensive review. *Neuroimage* **2014**, *85*, 72–91. [[CrossRef](#)] [[PubMed](#)]
46. Doğantekin, A.; Özyurt, F.; Avcı, E.; Koc, M. A novel approach for liver image classification: PH-C-ELM. *Measurement* **2019**, *137*, 332–338. [[CrossRef](#)]
47. Zafar, A.; Hong, K.-S. Detection and classification of three-class initial dips from prefrontal cortex. *Biomed. Opt. Express* **2017**, *8*, 367–383. [[CrossRef](#)] [[PubMed](#)]
48. Siero, J.C.; Hendrikse, J.; Hoogduin, H.; Petridou, N.; Luijten, P.; Donahue, M.J. Cortical depth dependence of the BOLD initial dip and poststimulus undershoot in human visual cortex at 7 Tesla. *Magn. Reson. Med.* **2015**, *73*, 2283–2295. [[CrossRef](#)] [[PubMed](#)]
49. Lu, H.D.; Chen, G.; Cai, J.; Roe, A.W. Intrinsic signal optical imaging of visual brain activity: Tracking of fast cortical dynamics. *Neuroimage* **2017**, *148*, 160–168. [[CrossRef](#)] [[PubMed](#)]
50. Akiyama, T.; Ohira, T.; Kawase, T.; Kato, T. TMS orientation for NIRS-functional motor mapping. *Brain Topogr.* **2006**, *19*, 1–9. [[CrossRef](#)]

51. Hu, X.; Yacoub, E. The story of the initial dip in fMRI. *Neuroimage* **2012**, *62*, 1103–1108. [[CrossRef](#)]
52. Wylie, G.R.; Graber, H.L.; Voelbel, G.T.; Kohl, A.D.; DeLuca, J.; Pei, Y.; Xu, Y.; Barbour, R.L. Using co-variations in the Hb signal to detect visual activation: A near infrared spectroscopic imaging study. *Neuroimage* **2009**, *47*, 473–481. [[CrossRef](#)]

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.