



Article

Inter-Subject MEG Decoding for Visual Information with Hybrid Gated Recurrent Network

Jingcong Li ^{1,2}, Jiahui Pan ^{1,2}, Fei Wang ^{1,2,*} and Zhuliang Yu ^{2,3}

- School of Software, South China Normal University, Guangzhou 510631, China; lijingcong@hotmail.com (J.L.); panjiahui@m.scnu.com (J.P.)
- ² Pazhou Lab, Guangzhou 510330, China; zlyu@scut.edu.cn
- Ollege of Automation Science and Engineering, South China University of Technology, Guangzhou 510006, China
- * Correspondence: scutauwf@foxmail.com

Abstract: As an effective brain signal recording technique for neuroscience, magnetoencephalography (MEG) is widely used in cognitive research. However, due to the low signal-to-noise ratio and the structural or functional variabilities of MEG signals between different subjects, conventional methods perform poorly in decoding human brain responds. Inspired by deep recurrent network for processing sequential data, we applied the gated recurrent units for MEG signals processing. In the paper, we proposed a hybrid gated recurrent network (HGRN) for inter-subject visual MEG decoding. Without the need of any information from test subjects, the HGRN effectively distinguished MEG signals evoked by different visual stimulations, face and scrambled face. In the leave-one-out cross-validation experiments on sixteen subjects, our method achieved better performance than many existing methods. For more in-depth analysis, HGRN can be utilized to extract spatial features and temporal features of MEG signals. These features conformed to the previous cognitive studies which demonstrated the practicality of our method for MEG signal processing. Consequently, the proposed model can be considered as a new tool for decoding and analyzing brain MEG signal, which is significant for visual cognitive research in neuroscience.

Keywords: inter-subject; brain signal decoding; spatial/temporal features



Citation: Li, J.; Pan, J.; Wang, F.; Yu, Z. Inter-Subject MEG Decoding for Visual Information with HGRN. *Appl. Sci.* **2021**, *11*, 1215. https:// doi.org/10.3390/app11031215

Academic Editor: Jing Jin Received: 4 January 2021 Accepted: 26 January 2021 Published: 28 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Human brain has an excellent visual system that the main content of a scene can be captured in less than one second. As an important research aspect in neuroscience, it is significant to record brain signals and decode visual information from these signals [1]. With brain decoding techniques, we are able to understand the high-level recognition abilities and diagnose the psychological illness. Magnetoencephalography (MEG) is an advanced technique for recording brain magnetic signals within hundreds of channels. Based on the MEG technique, neuroscientists are able to study brain function in-depth [2–5]. Moreover, the sampling rate of MEG signal can be quite high that it can record brain magnetic signals in milliseconds. With high temporal resolution, MEG signals can be used to study the dynamic changes of brain function [6–8].

However, due to the low signal-to-noise ratio and the structural or functional variability of MEG signals between different subjects, conventional machine learning methods perform poorly in decoding MEG signals. In addition, the existing MEG decoding techniques were rare and had never been evaluated on common benchmark database. In the existing publications, MEG decoding methods were evaluated by different databases that the experimental performance varied a lot. As a result, the corresponding results maybe not quite convincing. Therefore, to facilitate the development of MEG decoding techniques, a MEG decoding experimental dataset was released [4,8].

Appl. Sci. 2021, 11, 1215 2 of 12

The experimental procedure of this dataset is shown in Figure 1. During the experiments, the subject was asked to look at pictures with face image or scrambled face image. In the meanwhile, a MEG recording system will record his/her brain magnetic signals. As shown in Figure 1, we can find that the MEG signals in different scalp locations had different signal components. Based on the recorded MEG signals and the corresponding visual stimulations, a decoding model can be built to predict the visual information (face or scrambled face image) from MEG signals. Consequently, different MEG decoding methods can be trained and evaluated by a common dataset [4,8].

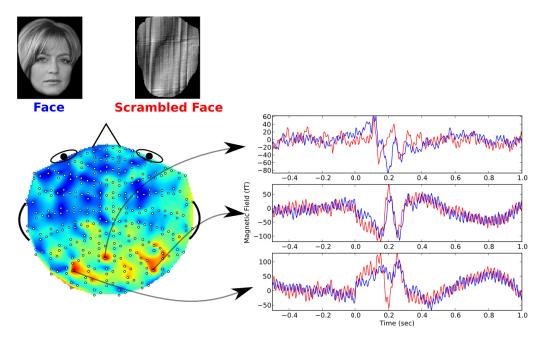


Figure 1. Magnetoencephalography (MEG) decoding for visual information experiment [4,8].

According to previous research, conventional methods usually rely on ensemble averaging technique to deal with the low signal-to-noise ratio of brain signals and acquire reliable detection [9]. However, the ensemble averaging technique might erase some important information in each trial of ERP signals [10]. According to some neuroscience researches, continuous gamma effect was found in the averaged brain signals in multiple trials [11,12]. It means that studying single-trial signals is essential for understanding the dynamic change of human brain. Moreover, decoding brain signals in an inter-subject manner is significant in neuroscience researches [4,13,14]. For example, a transductive transfer learning (TTL) algorithm was proposed for inter-subject MEG decoding with single-trial samples [4]. To extract manifold features of MEG signals, a new MEG decoding method based on riemannian metric was applied, which achieved a high decoding accuracy [15]. Moreover, as a popular machine learning approach in many areas, random forest was applied for MEG decoding [16]. The random forest algorithm can effectively extract MEG features of different subjects and improve inter-subject MEG decoding performance.

In the past few years, deep learning methodology was widely used and achieved remarkable results in many areas. As for brain signal processing, deep learning methods also worked well in extracting features, detecting abnormal components, classifications and so on. For example, deep network based on restricted Boltzmann machine (RBM) achieved good performance in processing electroencephalogram(EEG) signals [17,18]. Convolutional neural network (CNN) was firstly applied for detecting event-related potentials from EEG signals [19]. Moreover, deep learning method was utilized to study the dynamics of scene representations in the human brain revealed by MEG [20]. To analyze the MEG signals with high temporal resolution, a CNN-based model was proposed that the experimental results demonstrated its efficiency [21]. Logistic regression with ℓ 1 penalization (termed Pool) was proposed to model MEG data of many subjects [4,22,23]. Inspired by deep learning

Appl. Sci. 2021, 11, 1215 3 of 12

methodology, two logistic regression layers were stacked that a new MEG decoding method termed stacked generation (SG) was proposed and achieved better performance than the Pool model [4,24]. Additionally, inspired by deep learning models, logistic regression and random forest were combined in a deep structure (called linear regression and random forest—LRRF model) which achieved a good MEG decoding performance [25]. A deep network based on riemannian metric was proposed for decoding MEG signals evoked by the stimulation of face or scrambled face images [15,26].

However, there are still many problems in the existing MEG decoding methods. We need to find a more effective approach to model numerous high-dimensional MEG signals, extract their features and further improve MEG decoding performance. The main contributions of this paper can be summarized as follows:

- 1. A hybrid gated recurrent network (HGRN) was proposed for inter-subject visual MEG decoding.
- 2. The proposed HGRN is able to model high-dimensional MEG signal, extract its features, and distinguish the signal evoked by different visual stimulations.
- 3. Experimental results demonstrated that the proposed model is able to achieve a good MEG decoding performance. More importantly, the proposed method can be utilized as a new analyzing tool for MEG signals.

The remainder of this paper is organized as follows. The proposed hybrid gated recurrent network (HGRN) is presented in Section 2. In Section 3, numerical inter-subject MEG decoding experiments are carried out. In addition, the performance of different MEG decoding methods are presented and compared. Some discussions are presented in Section 4. Conclusions of this paper are given in Section 5.

2. Proposed Method

For many conventional methods, a multi-channel MEG sample with spatial-temporal structure must be reshaped as a feature vector that its original spatial/temporal structure will be corrupted. This maybe one of the main reasons for their low performance. To deal with the spatial-temporal structure of MEG signal, we proposed a hybrid gated recurrent network (HGRN) which is a modified recurrent network based on gate recurrent units (GRU). To our best knowledge, the HGRN is the first gated recurrent network for MEG decoding. Compared with the conventional methods, the HGRN can directly process multichannel MEG signal, extract its spatial-temporal features and decode it within one network.

The gated recurrent unit (GRU) is a special network structure which utilizes multiple gating units to store or regulate information [27–29]. In a GRU structure, its gated units are only activated by the current inputs and the previous inputs. Compared with the other kind of recurrent network structure, the GRU consists of less parameters and is converged faster as well. The GRU can adaptively capture the dependencies of the input data on different time scales. Many researches showed that the deep networks based on GRU achieved high performance in the sequence processing tasks [30–32].

In Figure 2, the internal structure of GRU is shown. The output of GRU at time t is $\mathbf{h}_t = \{h_t^1, ..., h_t^j, ...\}$ where j denotes the j-th element. The h_t^j can be obtained by linear interpolation between the previous output h_{t-1}^j and the current candidate output \tilde{h}_t^j as follow:

$$h_t^j = (1 - z_t^j) h_{t-1}^j + z_t^j \tilde{h}_t^j \tag{1}$$

where z_t^j is the j element of update gate \mathbf{z}_t . The z_t^j determines the updated scale of the output of GRU.

Appl. Sci. 2021, 11, 1215 4 of 12

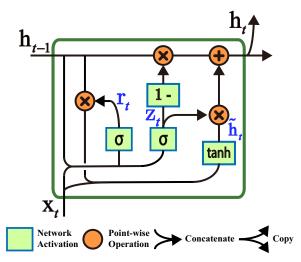


Figure 2. Gated recurrent unit (GRU).

Here, the update gate z_t^j can be obtained by the following statement:

$$z_t^j = \sigma(W_z \mathbf{x}_t + U_z \mathbf{h}_{t-1})^j \tag{2}$$

where \mathbf{x}_t denotes the input value at time t, $\mathbf{h}_{t-1} = \{h_{t-1}^1, ..., h_{t-1}^j, ...\}$ is the previous output at time t-1.

As the traditional recurrent unit, the candidate output \tilde{h}_t^j is computed by:

$$\tilde{h}_t^j = \tanh \left(W \mathbf{x}_t + U(\mathbf{r}_t \odot \mathbf{h}_{t-1}) \right)^j \tag{3}$$

where $\mathbf{r}_t = \{r_t^1, ..., r_t^J, ...\}$ is the reset gate, \odot is the Hadamard product (element-wise product), the tanh(\cdot) is a nonlinear function which is computed by

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4}$$

Following the update gate, the reset gate \mathbf{r}_t is defined as

$$r_t^j = \sigma(W_r \mathbf{x}_t + U_r \mathbf{h}_{t-1})^j \tag{5}$$

When the reset gate is closed, i.e., $r_t^j=0$, the GRU will forget the previous states, and read the current input sequence instead. Based on the update gate z_t^j and the reset gate r_t^j , the GRU can remember and update its states which is helpful for adaptively capturing the dependencies of input sequences in different time scales.

Based on the GRU structure, the proposed HGRN is built to adaptively capture the dependencies of MEG signals in different time scales. In Figure 3, the structure of the proposed HGRN is shown. The proposed HGRN consists of two GRU-based layers, two down-sampling layers and an output layer. In the GRU-based layers, there are two GRUs, the GRU1 and GRU2 layer. Located in two different layers, GRU1 and GRU2 consist of 100 and 10 output units, respectively. The GRU layers are able to extract spatial features at the current time point and the previous ones. Compared with the other kinds of network structures, the HGRN model with GRU layer can efficiently process spatial-temporal data with a simple structure. In the down-sampling layers, max pooling (Max-Pool) operation is applied [33]. The Max-Pool(5) indicates that the max activation value at 5 contiguous time points is considered as the sampled value. Then the temporal features within 20 ms and 100 ms scale can be extracted by the first and second max pooling layers. Moreover, the output layer of HGRN is a fully connected layer with two output units. The spatial and temporal features can be combined to achieve better MEG decoding performance. Similar

Appl. Sci. 2021, 11, 1215 5 of 12

to many other deep networks, the HGRN applies softmax function for the activations of its output units. Then, the outputs of HGRN can be obtained that the visual stimulation of the current input MEG signal is predicted.

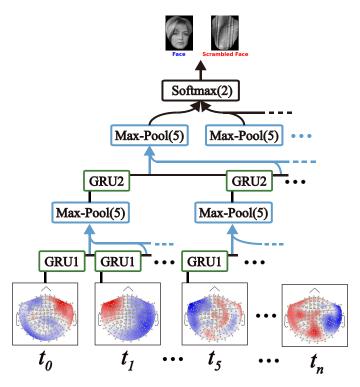


Figure 3. Hybrid gated recurrent network (HGRN).

Accordingly, the HGRN model can be optimized by minimizing the cross-entropy between its predictions and the true labels. Therefore, the corresponding loss function of the proposed model is defined as follow:

$$L = -\sum_{c=1}^{2} y_c \log(p_c) + (1 - y_c) \log(1 - p_c)$$
 (6)

where c = 1 denotes the MEG signal evoked by face image, c = 2 denotes the MEG signal evoked by scrambled face image, p_1 is the probability of the HGRN model to detect a MEG signal evoked by face image, p_2 is the probability corresponding to scrambled face image stimulation, $[y_1, y_2]^T$ indicate the ground truth of the input signal. If the input MEG signal is stimulated by face image, the ground truth is $[y_1, y_2]^T = [1, 0]^T$, otherwise, $[y_1, y_2]^T = [0, 1]^T$.

In the next section, a series of experiments will be carried out to evaluate the proposed HGRN model. In addition, the corresponding experimental results of our method will be presented and compared with the other methods.

3. Experiments

In this section, the experimental procedure and results of the proposed HGRN will be presented. In the experiments, the proposed method and some other popular MEG decoding methods are tested by a common dataset. For ease of reproduction, the details of data preprocessing and normalization will be illustrated. Moreover, the training procedure and the parameter settings of the proposed model are given as well.

In our experiments, the hardware and software configuration of our system is a desktop with Intel i7 3.6 Ghz CPU, 16 G DDR3 RAM, Nvidia Titan X, Ubuntu 16.04 and Keras with tensorflow backend.

Appl. Sci. **2021**, 11, 1215 6 of 12

3.1. MEG Signal Preprocessing

According to the experiments of the public visual MEG decoding dataset [4,8], the subject will stimulated by two different visual stimulations, face and scrambled face images. Sixteen subjects took part in the MEG decoding experiment. During the experiment, each subject received about 580 visual stimulations. The corresponding MEG signals were recorded by a MEG equipment with 306 channels. The provided MEG signals were downsampled to a rate of 250 Hz. Before each stimulation, the subject can rest 0.5 s and then receive a stimulation which will last one second. Firstly, the MEG signals evoked by each stimulation will be extracted. Secondly, each MEG signal was filtered by a [0.1, 20] Hz 4-order Butterworth filter. Lastly, each signal was normalized by subtracting its mean, and then dividing its standard deviation.

Consequently, the MEG data of all 16 subjects will be applied for leave-one-out cross-validation (LOOCV) in the following experiment.

3.2. Experimental Procedure

In order to train the proposed HGRN model, the Adam optimizer [34] was applied for minimizing the model's loss function (as shown in (6)). In each round of LOOCV experiment, the MEG data of 15 subjects were selected as training database, and the remaining one subject's MEG data were considered as validation database. During the training procedure, 90 percent of the samples in training database was randomly chosen as the training set while the remaining 10 percent was chosen as the observation set for monitoring the training procedure. Then the proposed HGRN model was trained by Adam optimizer with a learning rate of 1.0×10^{-3} , a learning rate decay parameter of 1.0×10^{-6} , a drop-out rate of 0.1 and mini-bath size of 100. The drop-out operation was only applied in the training procedure for randomly blocking the output of GRUs. During the training procedure, the model loss on the observation set was monitored. As the monitored loss was minimized, the training procedure was stopped.

3.3. Experimental Result

Once the proposed HGRN model was trained, it could be tested by the validation database that its MEG decoding performance can be evaluated. For the public MEG decoding database of 16 subjects, the LOOCV experiments were carried out 16 times so that the corresponding results could be considered as the performance references of different MEG decoding methods.

In Table 1, the results of the proposed HGRN model are presented. In addition, the experimental results of some state-of-the-art MEG decoding methods are also presented for comparison. To be clear, these methods are all supervised learning approaches which do not require any information of the testing data. Although some semi-supervised learning methods may achieve higher accuracy, they just utilized the information of the testing data which is not quite reasonable in practice [4]. In the method termed Pool [4,22,23], logistic regression with $\ell 1$ penalization was used to model the MEG data of multiple subjects, and applied for decoding MEG signals of new subject. The Pool method did not consider the differences between different subjects that the corresponding inter-subject MEG decoding performance was not very well. By applying two logistic regression layers, the stacked generalization (SG) model achieved better performance than the Pool method [4,24]. Considering the differences between subjects, stacked generalization with covariate shift (SG+CS) was proposed that different weighted parameters were applied for different subjects in the seconde logistic layer of SG+CS model [4,35]. Moreover, the MEG decoding method based on Riemannian geometry achieved a good performance [15,26]. To imitate deep structure of deep network, a method termed linear regression and random forest (LRRF) was proposed. The first layer of LRRF consisted of multiple logistic regression classifiers, and the output probabilities of the first layer was fed into a random forest for MEG decoding [25].

Appl. Sci. **2021**, 11, 1215 7 of 12

Sub.	Pool [22]	SG [36]	SG+CS [4]	Gen [15]	LRRF [25]	HGRN Ours
1	0.62	0.67	0.71	0.755	0.764	0.759
2	0.64	0.63	0.65	0.675	0.691	0.672
3	0.60	0.59	0.61	0.610	0.640	0.616
4	0.70	0.75	0.72	0.759	0.788	0.800
5	0.58	0.63	0.69	0.684	0.691	0.701
6	0.65	0.60	0.60	0.676	0.617	0.692
7	0.61	0.68	0.72	0.697	0.755	0.704
8	0.64	0.66	0.71	0.682	0.704	0.718
9	0.67	0.71	0.73	0.695	0.780	0.739
10	0.57	0.66	0.70	0.715	0.707	0.724
11	0.60	0.65	0.67	0.655	0.714	0.699
12	0.65	0.67	0.66	0.785	0.754	0.785
13	0.60	0.64	0.66	0.682	0.667	0.684
14	0.69	0.68	0.68	0.748	0.699	0.704
15	0.67	0.67	0.70	0.739	0.709	0.734
16	0.52	0.57	0.56	0.588	0.574	0.654
Mean	0.62	0.65	0.67	0.697	0.703	0.712

Table 1. Performance of different supervised learning methods for MEG decoding.

As shown in Table 1, the proposed HGRN achieved higher performance than many other methods in the LOOCV experiments of 16 subjects. Specifically, the averaged MEG decoding accuracy of HGRN was 0.712 which was the highest in the table. The experimental results demonstrated that the proposed HGRN is an effective model for inter-subject MEG decoding.

In order to compare the performance of the proposed method with the other methods, the Wilcoxon signed-rank test was applied for statistical analysis. As shown in Table 2, the Wilcoxon signed-rank test results between the proposed HGRN and the other methods are presented. Accordingly, the statistical analysis demonstrated that the HGRN achieved significantly better performance than the methods like Pool, SG, SG+CS and Gen. Although the proposed model did not achieve a significant better performance than LRRF, its averaged accuracy was a little higher than the LRRF as shown in Table 1. These results indicated the effectiveness of the proposed HGRN on MEG decoding performance.

Table 2. Wilcoxon signed-rank test between HGRN and other MEG decoding models.

Wilcoxon Signed-Rank Test	<i>p-</i> Value	
HGRN vs. Pool	++	
HGRN vs. SG	††	
HGRN vs. SG+CS	††	
HGRN vs. Gen	†	
HGRN vs. LRRF	~	

Note: \sim nonsignificant, † p < 0.05, †† p < 0.01.

In the next section, we will discuss and analyze the internal properties of the proposed model. It may be significant for further research.

4. Discussion

Based on GRU structure, the proposed HGRN model was built to capture the features of MEG signals and decode the signals. In this section, the proposed HGRN model will be discussed in details. First of all, the performance of the proposed model to decode face image stimulation and scrambled face image stimulation were presented in confusion matrix. Based on confusion matrix, the decoding performance difference between the two stimulations can be analyzed. Secondly, the weights of MEG sensors on the scalp for decoding signals are extracted. Consequently, the spatial distribution of the MEG signals

Appl. Sci. 2021, 11, 1215 8 of 12

in the decoding experiments can be determined. Lastly, the MEG signals evoked by two different stimulations will be extracted for analysis.

To analyze the MEG decoding performance of HGRN, confusion matrix is applied and presented in Figure 4. Specifically, the MEG signals evoked by face images were considered as positive class, while the MEG signals evoked by scrambled face images were categorized as the negative class. In the experiment, the true positive category indicated that the predicted results and the ground truth are all positive class. Likewise, the true negative category indicated that the predicted results and the ground truth are the same negative class. The false positive indicated that the predicted results are positive class while the ground truth is negative. The false negative indicated that the predicted results are negative class while the ground truth is positive. The TP, TN, FP and FN denote the number of samples in the categories of true positive, true negative, false positive and false negative, respectively. Then, the true positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR) are calculated. These results were calculated with all 16 subjects' results in the LOOCV experiments.

	Predicted label				
	Positive	Negative			
label Positive	TPR=70.93% (TP=3358)	FNR=28.53% (FN=1335)			
True Negative	FPR=29.07% (FP=1376)	TNR=71.47% (TN=3345)			

Figure 4. Confusion matrix of HGRN for all 16 subjects' decoding results.

As shown in Figure 4, the number of samples in the categories of true positive, true negative, false positive and false negative are presented in the confusion matrix. Moreover, the corresponding TPR, TNR, FPR and FNR are also presented. According to Figure 4, the TPR and TNR were closed to 71%. The corresponding recall, precision and F1 scores are 71%, 72% and 71%, respectively. These results demonstrated that the HGRN model can evenly distinguish the positive samples and negative samples (i.e., MEG signals evoked by face images and scrambled face images). The accuracy is a proper metric to evaluate balance datasets according to the previous study [37]. In the used dataset, the number of the positive and negative samples are 4693 and 4721 that it can be considered as a balance dataset. The accuracies presented in Table 1 can be used for the comparison between the performance of the other models and ours. In addition, about 30 percent of samples were incorrectly categorized. That is to say, there was still a lot of room for improving MEG decoding performance.

In order to study the spatial features of MEG signals, we applied a leave-one-alone strategy that only one MEG channel was reserved for analysis while the rest channels were clamped to zero. Next, the new MEG signals were tested by the HGRN model which is trained previously. Then, the MEG decoding result corresponding to each reserved channel can be obtained. To a certain extent, these results can be considered as the contributions of MEG channels in the decoding experiments. By normalizing these results to [0, 1], the weight of each channel for MEG decoding is calculated. Then, the spatial feature of each subject can be obtained. As shown in Figure 5, the spatial feature of MEG signals of Subject 1 is presented that its top view and side view are presented in the left subfigure and right subfigure, respectively. For ease of reading, the nose and ears of the subject were also presented for references in Figure 5. Likewise, the spatial features of MEG signals of all 16 subjects are obtained and presented in Figure 6. As shown in Figures 5 and 6, the MEG channel with the highest weight is usually located in the occipital region of each subject. This finding conformed to many previous research that the visual perception organization of the human brain is mainly located in the occipital region [7]. The experimental results and analysis indicate that the proposed HGRN model can be used for analyzing the

Appl. Sci. 2021, 11, 1215 9 of 12

spatial features of MEG signals. It shows a potential value of the proposed HGRN in the application of brain-computer interface and neuroscience research.

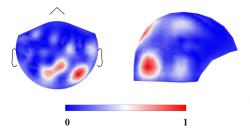


Figure 5. Spatial feature of MEG signals of Subject 1 (left: top view, right: side view).

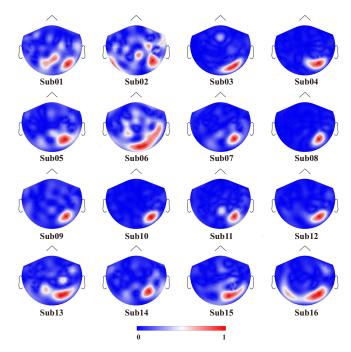


Figure 6. Spatial features of different subjects.

According to the spatial feature in Figure 6, the MEG channel with largest weight of each subject can be determined. To some certain extent, the signals at this channel can be considered as typical MEG signals for the decoding experiments. By averaging the MEG signals at the highest-weighted channel, the typical MEG waveforms evoked by face or scrambled face image stimulation are obtained. As shown in Figure 7, the typical MEG waveforms evoked by the two different stimulations are presented. From the figure, we can find that the waveforms evoked by face and scrambled face images were quite different at about 0.2 s. The waveforms corresponding to face image stimulation usually had lower amplitude at about 0.2 s according to the previous studies [4,13,14].

The extraction and analysis methods of the above MEG waveform can be applied in the research of neuroscience. For example, the change of brain cognitive activity over time can be explored. The proposed method can be considered as a new tool for further exploration of advanced human brain cognitive behavior.

Appl. Sci. 2021, 11, 1215

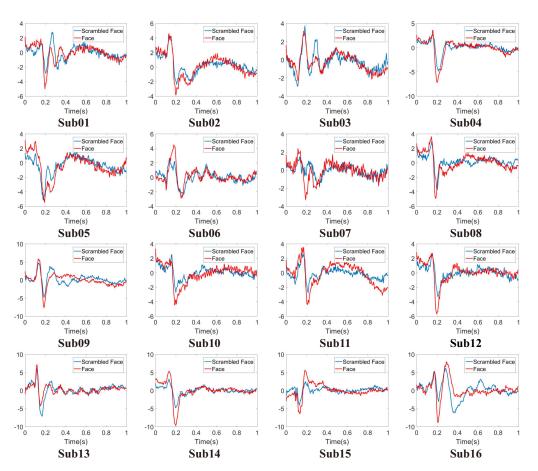


Figure 7. Typical MEG signals evoked by face or scrambled face stimulation.

5. Conclusions

In this paper, a hybrid gated recurrent network (HGRN) is proposed for inter-subject MEG decoding with visual information. Evoked by visual stimulations, the MEG signals were recorded by an equipment with hundreds of channels at a high sampling rate. By building a multi-layer recurrent structure, HGRN model can capture the features of MEG signals of different subjects. In the cross-validation experiment on a public database with 16 subjects, the proposed HGRN model achieved higher performance than many existing methods. Moreover, based on the HGRN, the spatial feature of MEG signals of each subject can be extracted for analysis. According to our experiment, we found that only the signals in occipital area play important roles in the decoding procedure which conformed to many previous studies. The above experimental results and analysis show that the proposed HGRN model can extract the features of MEG signals and effectively improve visual MEG decoding performance. In the future, we would like to build more efficient networks to model brain signals in multiple modalities. Moreover, some new emerging machine learning techniques can be a new inspiration for the methodology in brain-computer interface as well as neuroscience research.

Author Contributions: J.L. proposed the idea, conducted the experiments, and wrote the manuscript. J.P. provided advice on the research approaches and signal processing. F.W. checked and revised the manuscript. Z.L.Y. offered important help that guided the experiments and analysis methods. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under grant 62006082, 62076103 and 61836003, the Key Realm R and D Program of Guangzhou under grant 202007030005, the Guangdong Natural Science Foundation under grant 2020A1515110294, the SCNU Research Fund under grant 19KJ02, Guangdong Natural Science Foundation Doctoral Research Project (2018A030310365) and International Cooperation open Project of State Key Laboratory of Sub-

Appl. Sci. **2021**, 11, 1215

tropical Building Science, SCUT (2019ZA02), and the Guangdong General Colleges and Universities Special Projects in Key Areas of Artificial Intelligence of China under Grant 2019KZDZX1033.

Institutional Review Board Statement: The study mainly involves an open dataset and proposed a new algorithm to process the dataset that did not directly involve humans or animals.

Informed Consent Statement: The study did not directly involve humans.

Data Availability Statement: The used dataset is available at https://www.kaggle.com/c/decoding-the-human-brain and https://www.openfmri.org/dataset/ds000117/.

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

References

- 1. Thorpe, S.; Fize, D.; Marlot, C. Speed of processing in the human visual system. *Nature* 1996, 381, 520. [CrossRef] [PubMed]
- 2. Cecotti, H. Single-trial detection with magnetoencephalography during a dual-rapid serial visual presentation task. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 220–227. [CrossRef]
- 3. Abadi, M.K.; Subramanian, R.; Kia, S.M.; Avesani, P.; Patras, I.; Sebe, N. DECAF: MEG-based multimodal database for decoding affective physiological responses. *IEEE Trans. Affect. Comput.* **2015**, *6*, 209–222. [CrossRef]
- 4. Olivetti, E.; Kia, S.M.; Avesani, P. MEG decoding across subjects. In Proceedings of the 2014 International Workshop on Pattern Recognition in Neuroimaging, Tubingen, Germany, 4–6 June 2014; pp. 1–4.
- 5. Bradberry, T.J.; Rong, F.; Contreras-Vidal, J.L. Decoding center-out hand velocity from MEG signals during visuomotor adaptation. *Neuroimage* **2009**, *47*, 1691–1700. [CrossRef] [PubMed]
- 6. Hämäläinen, M.; Hari, R.; Ilmoniemi, R.J.; Knuutila, J.; Lounasmaa, O.V. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Modern Phys.* **1993**, *65*, 413.
- 7. Andersen, L.M.; Pedersen, M.N.; Sandberg, K.; Overgaard, M. Occipital MEG activity in the early time range (<300 ms) predicts graded changes in perceptual consciousness. *Cereb. Cortex* **2015**, *26*, 2677–2688.
- 8. Wakeman, D.G.; Henson, R.N. A multi-subject, multi-modal human neuroimaging dataset. Sci. Data 2015, 2, 150001. [CrossRef]
- 9. Luck, S.J. An Introduction to The Event-Related Potential Technique; MIT Press: Cambridge, MA, USA, 2005.
- 10. Stokes, M.; Spaak, E. The Importance of single-trial analyses in cognitive neuroscience. *Trends Cogn. Sci.* **2016**, 20, 483–486. [CrossRef]
- 11. Lowet, E.; Roberts, M.; Bosman, C.; Fries, P.; De Weerd, P. Areas V1 and V2 show microsaccade-related 3-4Hz covariation in gamma power and frequency. *Eur. J. Neurosci.* **2016**, 43, 1286–1296. [CrossRef]
- 12. Lundqvist, M.; Rose, J.; Herman, P.; Brincat, S.L.; Buschman, T.J.; Miller, E.K. Gamma and beta bursts underlie working memory. *Neuron* **2016**, *90*, 152–164. [CrossRef]
- 13. Bolagh, S.N.G.; Shamsollahi, M.B.; Jutten, C.; Congedo, M. Unsupervised cross-subject BCI learning and classification using Riemannian geometry. In Proceedings of the 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, 1 April 2016.
- 14. Westner, B.U.; Dalal, S.S.; Hanslmayr, S.; Staudigl, T. Across-subjects classification of stimulus modality from human MEG high frequency activity. *PLoS Comput. Biol.* **2018**, *14*, 1–14. [CrossRef] [PubMed]
- 15. Barachant, A. MEG Decoding Using Riemannian Geometry and Unsupervised Classification; Grenoble University: Grenoble, France, 2014.
- Fatima, S.; Kamboh, A.M. Decoding brain cognitive activity across subjects using multimodal M/EEG neuroimaging. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea, 11–15 July 2017; pp. 3224–3227.
- 17. Lu, N.; Li, T.; Ren, X.; Miao, H. A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines. *IEEE Trans. Rehab. Eng.* **2017**, 25, 566. [CrossRef] [PubMed]
- 18. Li, J.; Yu, Z.L.; Gu, Z.; Wu, W.; Li, Y.; Jin, L. A Hybrid Network for ERP Detection and Analysis Based on Restricted Boltzmann Machine. *IEEE Trans. Rehab. Eng.* **2018**, *26*, 563–572. [CrossRef] [PubMed]
- 19. Cecotti, H.; Graser, A. Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 433–445. [CrossRef] [PubMed]
- 20. Cichy, R.M.; Khosla, A.; Pantazis, D.; Oliva, A. Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. *Neuroimage* **2017**, *153*, 346–358. [CrossRef] [PubMed]
- 21. Seeliger, K.; Fritsche, M.; Güçlü, U.; Schoenmakers, S.; Schoffelen, J.M.; Bosch, S.; van Gerven, M. Convolutional neural network-based encoding and decoding of visual object recognition in space and time. *NeuroImage* **2018**, *180*, 253–266. [CrossRef]
- 22. Tibshirani, R. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Methodol.) 1996, 58, 267–288. [CrossRef]
- 23. Tomioka, R.; Müller, K.R. A regularized discriminative framework for EEG analysis with application to brain-computer interface. *NeuroImage* **2010**, *49*, 415–432. [CrossRef]
- 24. Wolpert, D.H. Stacked generalization. Neural Netw. 1992, 5, 241–259. [CrossRef]
- 25. Huttunen, H.; Gencoglu, O.; Lehmusvaara, J.; Vartiainen, T. MEG Decoding with Hierarchical Combination of Logistic Regression and Random Forests; Technical Report, DecMeg 2014 Competition; Tampere University of Technology: Tampere, Finland, 2014.

Appl. Sci. 2021, 11, 1215

26. Caliskan, A.; Yuksel, M.E.; Badem, H.; Basturk, A. A deep neural network classifier for decoding human brain activity based on magnetoencephalography. *Elektron. ir Elektrotechnika* **2017**, 23, 63–67. [CrossRef]

- 27. Cho, K.; Van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv* **2014**, arXiv:1409.1259.
- 28. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
- 29. Jozefowicz, R.; Zaremba, W.; Sutskever, I. An empirical exploration of recurrent network architectures. In Proceedings of the 32nd International Conference on Machine Learning Lille, France, 6–11 July 2015; pp. 2342–2350.
- Dwibedi, D.; Sermanet, P.; Tompson, J. Temporal Reasoning in Videos Using Convolutional Gated Recurrent Units. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Salt Lake City, UT, USA, 8–22 June 2018.
- 31. Nilsson, D.; Sminchisescu, C. Semantic Video Segmentation by Gated Recurrent Flow Propagation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 8–23 June 2018.
- 32. Jing, L.; Gulcehre, C.; Peurifoy, J.; Shen, Y.; Tegmark, M.; Soljacic, M.; Bengio, Y. Gated Orthogonal Recurrent Units: On Learning to Forget. *Neural Comput.* **2019**, *31*, 765–783. PMID: 30764742. [CrossRef] [PubMed]
- 33. Jaderberg, M.; Simonyan, K.; Zisserman, A.; Kavukcuoglu, K. Spatial transformer networks. arXiv 2015, arXiv:1506.02025.
- 34. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. arXiv 2014, arXiv:1412.6980.
- 35. Shimodaira, H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *J. Stat. Plan. Inference* **2000**, *90*, 227–244. [CrossRef]
- 36. Ting, K.M.; Witten, I.H. Issues in stacked generalization. J. Artif. Intell. Res. 1999, 10, 271–289. [CrossRef]
- 37. He, H.; Garcia, E.A. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 2009, 21, 1263-1284.