Motor Imagery Electroencephalogram Datasets, Where and How to Find Them: A Review

Z.M. Alhakeem[1,a]**, Heba Hakim**[2,b]**, and Ola.A. Hassan**[3,c,*]

1 Communications Engineering Department, Iraq University College, Basrah - Iraq.
2 Computer Engineering Department, University of Basrah, Basrah - Iraq.
3 Communications Engineering Department, Iraq University College, Basrah - Iraq.
E-mail: [a]zaineb.alhakeem@iuc.edu.iq , [b,*]hebah.hakem@gmail.com , [c,*]ola.ahmed@iuc.edu.iq

Abstract. In some research fields finding the suitable data to be used in testing and validating the methods is a very important step. Sometimes the researcher cannot find the exact dataset that is suitable to the research field, in this paper a review about how to find or create a dataset of electroencephalogram (EEG) signals in Motor imagery applications .

Keywords: Dataset, EEG, Motor Imagery, Recording device, Recording protocol.

1. Introduction

Finding the datasets in research is very important step, usually the researchers depend on the published datasets that are recorded and collected from well-known institutions that has the same specialty in the research field. EEG signals analysis field researches depend on testing the brain signals in different areas of the brain, these brain areas are shown in Fig. 1 [1]. The brain areas are located by standards that is known as (10/20) as shown in Fig. 2 [2][3]. The application also specifies the type of data and from which area should be recorded. The studies that used EEG signals have many varieties such as: sleep studies, epilepsy seizures diagnosis and prediction, emotions classification, Brain Computer Interface (BCI), ..etc [4]. The published data set are usually used to test and validate the method before implement them practically. Usually, the published data are good to test the methods in a controlled environment, but when the work should be generalized the data should be recorded in real-time. Besides that, in practical implementations of such studies such as navigating wheelchair using EEG signals or such studies the previously recorded data are not very useful, so the researcher should use his/her own recording device and data to implement the work such as in [5]. In this paper a review about how to record EEG data and how to find them if they are published already.
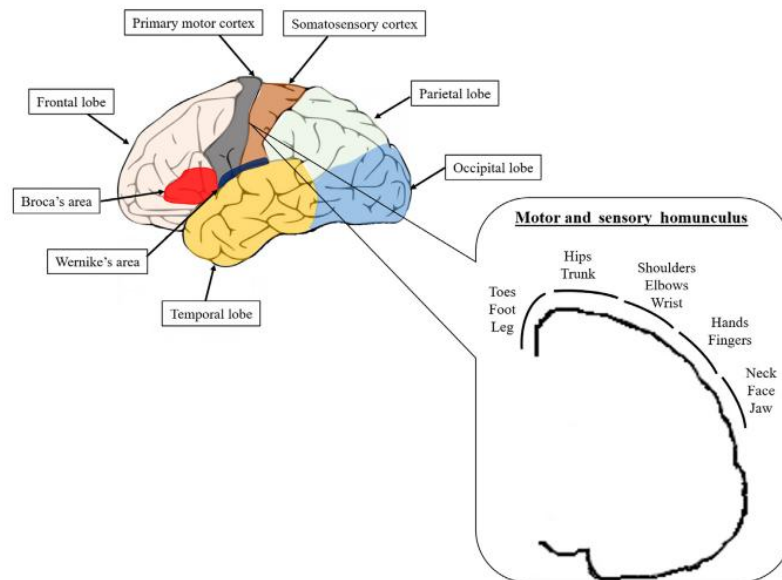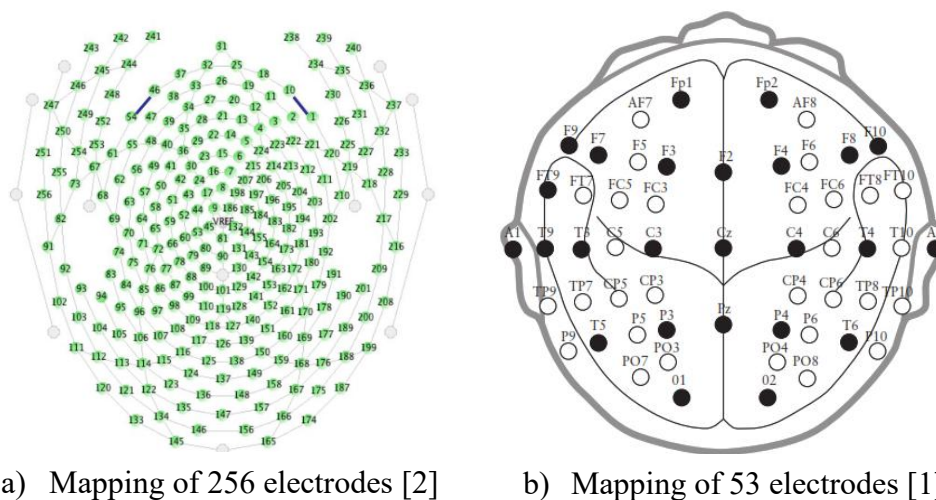

Fig. 1. The brain areas [1]

a) Mapping of 256 electrodes [2] b) Mapping of 53 electrodes [1]

Fig. 2. The electrodes distribution on brain areas according to 10/20 standard.

2. Protocols and Devices of Recording EEG Motor Imagery Signals

First of all, to understand the EEG datasets and how to find the suitable one for the work one should understand and have a knowledge about the devices that record such data and their prices. This step is important to evaluate whether the work could be practically implemented later or not. Sometimes the recording devices are very expensive so the application even if it is very useful and novel but its implementation is far from reality because of economic reasons or the recording device is not portable and the application require mobile device. Secondly the recording protocol should be clear to both the researcher and the people in the experiment. In this section the recording protocol and the recording devices are discussed.

2.1 Recording devices of EEG signals

There are many types of EEG data recording devices and sensors, depending on the application. EEG signals can be recorded using two mainly types of sensors. The first type is the invasive type which means that, recording electrodes are planted inside the brain, this type is out of the scope of this paper because it needs surgical procedures which is dangerous to apply by people that are not professional. The second type is the noninvasive sensors which is a type that attached to the skin without any other procedures, this type can be further divided into two types, dry and wet types.

The dry sensors are attached directly to the skin with no need to any other materials to be added for any reason while the wet one need some type of gels or alcohol to wet the electrodes and decrease the impedance of the skin to the signal. The dry sensors have an advantage of no need to buy the extra materials such as gel or alcohol periodically, no need to maintenance, but the disadvantage is the signal is not clear as the wet one. Table 1 shows the available portable recording devices of EEG signals. All the sensors in the Table 1 are need some kind of gel or silent to reduce the impedance of the skin except OpenBCI which has spiky sensors that not need any material to be between the skin and the sensor.

Table 1. EEG signals portable recording devices [6]

Recording Device	Number of electrodes	Sensors	Price
Emotiv EPOC X	14	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4	799$
Emotiv EPOC+	14	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4	699$
Emotiv INSIGHT	5	AF3,AF4,T7,T8,Pz	299$
Emotiv EPOC FLEX	Up to 32	FP1,FP2,F7,F8,F3,F4,Fz,T7,T8,P3,P4, P7,P8,Pz,C3,C4,Cz,O1,O2,Oz,CP3,CP 4,CP5,Cp6,FC5,FC6,FC1,FC2,FT9,FT 10,Tp10,Tp9	1699$
Open BCI	Up to 16	FP1,FP2,F7,F8,F3,F4,T7,T8,P3,P4,P7, P8,Pz,C3,C4,O1,O2,Oz	Up to1080$
NeuroSky MindWave	1	FP1	199$
g.SCARABEO	Up to 64	Cover Wide area of the head	Up to 35,990$
NeuroScan	Up to 256	Cover very wide area of the head	Depends on the type of amplifier

2.2 Recording Procedure of EEG Motor Imagery Signals

Usually, the recording procedure of EEG signal starting with preparing the participants of the experiments to the recording procedure taking the agreements from them and follow the instructions to record the suitable data. The participants will be seated comfortably on armed chair to prevent any limbs movements during data recording that will appeared as artifacts, they seated in front of a screen that will show the instructions on a Graphical User Interface (GUI) window. Fig. 3 shows the GUI of recording motor imagery EEG signal in which the instructions are written in Arabic language and there is an arrow pointed to the direction of the hand to be imagined. This GUI instruct the participants to imagine the movement of the Left hand and the Right hand to record the EEG signals of these movements [7]. Other GUIs are shown only a picture or a shape to instruct the participant to do the task such as shown in Fig.4 [8]. The instructions appear for 5 seconds, some other GUI view the instructions for ten seconds, as a rule the instructions should be appeared for enough time to the participant to understands the task and do it during the time of device recording.

Fig. 3. Graphical User Interface (GUI) of Left/Right hands motor imagery tasks using both words and shapes [7]
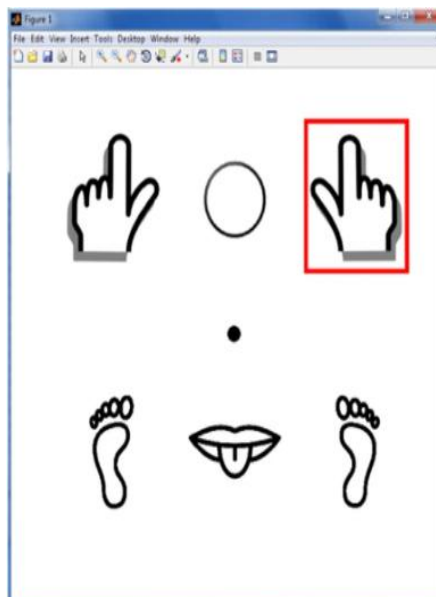

Fig. 4. Graphical User Interface (GUI) of Left/Right hands and feet with tongue using only images [8]

3. Published Datasets of EEG Motor Imagery Signals

The published datasets are very useful when the researcher want only to test the analysis and methods of signal processing without the need to implement them practically. These datasets are low cost or free for academic usage because the researcher does not need to buy very cost device to collect data. These datasets usually recorded in Institutions Labe rotary under supervised conditions for research reasons. Few time the researchers record their own dataset if they have available device and they publish the set to help other researchers in the same field.

3.1. BCI Competitions

Berlin Brain-Computer Interface is a research program that aims to study and develop Brain computer Interface (BCI) decoding and detection methods. This project record different types of data and make a competition among the researchers to achieve new contribution to this field. There are four published datasets each one according to a competition.

Unfortunately, the data of the first one is not available now on the web, but the rest three are still available on [9] and used very widely in the publications such as [10][11][12][13][14] and many other publications. This dataset has multiple sub datasets that contains data recorded for single user, three users, five users and ten users. Number of classes that available are two classes using two recording electrodes, two classes using three recording electrodes, two classes using 64 electrodes, two classes using 118 recording electrodes, four classes using 22 electrodes, four classes using 60 electrodes.

3.2. Dr. Cichocki's Lab dataset

This dataset is very useful when the researcher needs to test multiple session data for the same user or multiple user data for the same task. This dataset is recorded using two recording devices which they are (g.tec and Neuroscan). Eight users are recording this dataset using five, six and fourteen electrodes. Two and three classes are recorded for each user, this dataset is used to test the methods of the works of [15][10][16] and other publications.

4. Conclusion

In some research fields data are required to test and validate the methods that are contributed by the researcher. The brain signals research field not an exception, this field need EEG signals to be used. One type of these signals is Motor Imagery EEG signal, this type of signals needs some special treatment in recording environment and devices. The devices different in price, number of channels and signal resolution. The selection of the suitable recording device depends on the application. As example the practical applications for ordinary users cannot use a device with very high price because no one will buy such application. So, using the published datasets which recorded under supervision and with very high resolution could not give the real accuracy of the method but it is very useful to find out whether the method is successful or not. Some devices could not record the correct signal because the device does not have the sensor in the correct place, i.e. if we need to record a motor imagery signal it is better to record it using C3,C4 and the sensors around the motor cortex area of the brain some devices has no such sensor in this area, so the device is not suitable for the application not a bad one.

References

[1] P. Diez, *Smart Wheelchairs and Brain-Computer Interfaces*. Elsevier, 2018.

[2] W. O. Tatum *et al.*, *Handbook of EEG Interpretation*, 2nd ed. New York: Springer Publishing Company.

[3] "https://www.mamem.eu/results/datasets/," *dataset*. .

[4] L. Luis and J. Gomez-Gil, "Brain Computer Interfaces, a Review," *Sensors (Basel).*, vol. 12, pp. 1211–1279, 2012, doi: 10.3390/s120201211.

[5]   Z. M. Alhakeem, R. S. Ali, and R. A. Abd-Alhameed, "Wheelchair Free Hands Navigation Using Robust DWT-AR Features Extraction Method with Muscle Brain Signals," *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2984538.

[6]   D. Dadebayev, W. W. Goh, and E. X. Tan, "EEG-based emotion recognition: Review of commercial EEG devices and machine learning techniques," *J. King Saud Univ. - Comput. Inf. Sci.*, 2021, doi: https://doi.org/10.1016/j.jksuci.2021.03.009.

[7]   Zaineb M. Alhakeem, "Design and Implementation of Brain Computer Interface Systems," 2020.

[8]   M. Kaya, M. K. Binli, E. Ozbay, H. Yanar, and Y. Mishchenko, "A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces," *Sci. Data*, vol. 5, no. 1, p. 180211, 2018, doi: 10.1038/sdata.2018.211.

[9]   EMOTIV websit,"https://www.emotiv.com/." .

[10] R. S. A. Zaineb M. Alhakeem, "Session to Session Transfer Learning Method Using Independent Component Analysis with Regularized Common Spatial Patterns for EEG-MI Signals," *Iraqi J. Electr. Electron. Eng.*, vol. 15, no. 1, pp. 13–27, 2019.

[11] D. Lee, S.-H. Park, and S.-G. Lee, "Improving the Accuracy and Training Speed of Motor Imagery Brain-Computer Interfaces Using Wavelet-Based Combined Feature Vectors and Gaussian Mixture Model-Supervectors.," *Sensors (Basel).*, vol. 17, no. 10, Oct. 2017, doi: 10.3390/s17102282.

[12] A. Liu, K. Chen, Q. Liu, Q. Ai, Y. Xie, and A. Chen, "Feature Selection for Motor Imagery EEG Classification Based on Firefly Algorithm and Learning Automata.," *Sensors (Basel).*, vol. 17, no. 11, Nov. 2017, doi: 10.3390/s17112576.

[13] X. Wang *et al.*, "A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization From Chest CT," *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020, doi: 10.1109/TMI.2020.2995965.

[14] C. Yaacoub, G. Mhanna, and S. Rihana, "A Genetic-Based Feature Selection Approach in the Identification of Left/Right Hand Motor Imagery for a Brain-Computer Interface.," *Brain Sci.*, vol. 7, no. 1, Jan. 2017, doi: 10.3390/brainsci7010012.

[15] Z. M. Alhakeem and R. S. Ali, "Fast channel selection method using crow search algorithm," 2019, doi: 10.1145/3321289.3321309.

[16] D. C. V. J. Anila Maily and D. M. M. Sathikd, "Feature Extraction of Motor Imagery EEG Data Using Time Domain Statistical Parameters," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 10, pp. 483–490, 2021.

Zaineb M. Alhakeem was born in Basrah, Iraq in 1983. She received the B.S. and M.S. degrees in Computer engineering from the University of Basrah, in 2005 and 2008 respectively. She is also had her Ph.D degree within the Electrical Engineering Department, College of Engineering in 2020, University of Basrah, Basrah, Iraq.

Her teaching interests covering wide areas of modules across the department of Computer Engineering and Communication Engineering, Iraq University College, include Neural Networks, Control Systems, Microprocessor & Microcontrollers, Programming Languages. Her research interests include Biomedical Engineering, Digital signal processing. Dr. Alhakeem is a Member in the IEEE.

Heba Hakim was born in Basrah, Iraq in 1980. She received the B.S. and M.S. degrees in Computer Engineering Department from the University of Basrah, in 2002 and 2004 respectively. She is also had the Ph.D. degree in Electrical Engineering Department in 2020 from University of Basrah, Basrah, Iraq.

Her research interests include Artificial Intelligence, Navigation System, Image Encryption and Steganography.
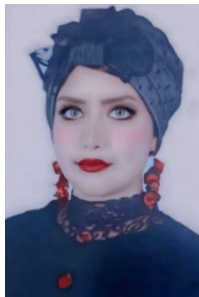
Ola A. Hasan was born in Iraq. She received the B.S. degree in computer engineering from the University of Basra in 2013 and the M.S. degree in electrical engineering also from the University of Basra in 2017.

In 2013, she worked as an IT engineer in the Basra Governorate Building. Since 2018, she has worked as a lecturer at Iraq University College. Her research interests include robotics, wireless sensor networks, biometrics, renewable energy, control systems, IoT, and programming. She is a member of the reviewing committees of a number of journals and international conferences.

She is a member of both the IEEE and ACM organizations..

مجلة كلية العراق الجامعة للهندسة والعلوم التطبيقية

قواعد بيانات إشارات الدماغ الكهربائية لتخيل حركات الأطراف، اين و كيف تجدها :مراجعة

زينب محمد الحكيم[1]،[أ] ، هيه حاكم عبدالزهرة [2]،[ب]، علا احمد حسن[3]،[ج]

1كلية العراق الجامعة , قسم هندسة الاتصالات , البصرة , العراق.
2جامعة البصرة, كلية الهندسة, قسم هندسة الحاسبات , البصرة , العراق.

البريد الالكتروني : [أ]zaineb.alhakeem@iuc.edu.iq[ب]hebah.hakem@gmail.com، [ج]ola.ahmed@iuc.edu.iq,

الملخص. في بعض مجالات البحث نحتاج لقاعدة البيانات لتطبيق الطرق التي تم التوصل لها اثناء البحث هي خطوة مهمة جدا. في بعض الأحيان البحاث غير قادر على إيجاد قاعدة البيانات المناسبة . في هذا البحث مراجعة لقواعد البيانات المتوفرة و أجهزة تسجيل البيانات و أنواعها لاشارات الدماغ المتولدة من تخيل حركة الأطراف.