Data in Brief 40 (2022) 107675

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Music genre neuroimaging dataset

Tomoya Nakai[a,b,c], Naoko Koide-Majima[a,d], Shinji Nishimoto[a,d,e,*]

[a] Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita, Japan
[b] Lyon Neuroscience Research Center (CRNL), INSERM U1028 - CNRS UMR5292, University of Lyon, Lyon, France
[c] JSPS Overseas Research Fellow, Tokyo, Japan
[d] Graduate School of Frontier Biosciences, Osaka University, Suita, Japan
[e] Graduate School of Medicine, Osaka University, Suita, Japan

ARTICLE INFO

Article history:
Received 31 August 2021
Revised 25 November 2021
Accepted 1 December 2021
Available online 4 December 2021

Keywords:
Music genre recognition
Brain activity
fMRI
Machine learning

ABSTRACT

This dataset includes functional magnetic resonance imaging (fMRI) data collected while five subjects extensively listened to 540 music pieces from 10 music genres over the course of 3 days. Behavioral data are also available. Data are separated into training and test samples to facilitate the application of machine learning algorithms. Test stimuli were repeated four times and can be used to evaluate the signal to noise ratio of brain activity. Using this dataset, both neuroimaging and machine learning researchers can test multiple algorithms regarding the prediction performance of brain activity induced by various music stimuli. Although two previous studies have used this dataset, there remains much room to apply different acoustic models. This dataset can contribute to integration of the fields of auditory neuroscience and machine learning.

* Corresponding author
E-mail address: nishimoto@fbs.osaka-u.ac.jp (S. Nishimoto).
Social media: (T. Nakai), (S. Nishimoto)

https://doi.org/10.1016/j.dib.2021.107675

Specifications Table

Subject	Neuroscience: Cognitive
Specific subject area	Neuroimaging of human subjects listening to music stimuli from 10 different genres
Type of data	Images
How data were acquired	3T Siemens MRI (TIM Trio), 32-channel head coil. Presentation software
Data format	Raw
Parameters for data collection	All subjects are right-handed, with normal hearing, normal or corrected to normal vision, and no history of psychological or neurological disorders.
Description of data collection	Five subjects underwent a 3-day MRI experiment and a one-day behavioral experiment. Subjects listened to 540 music pieces from 10 different music genres.
Data source location	Institution: Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT) City/Town/Region: Osaka Country: Japan
Data accessibility	(1) Repository name: OpenNeuro, Data identification number: 10.18112/openneuro.ds003720.v1.0.0, Direct URL to data: https://openneuro.org/datasets/ds003720/versions/1.0.0 (2) Repository name: OSF, Data identification number: 10.17605/OSF.IO/2ZWKT, Direct URL to data: https://osf.io/2zwkt/
Related research article	(1) Nakai, T., Koide-Majima, N., & Nishimoto, S. (2021). Correspondence of categorical and feature-based representations of music in the human brain. *Brain and behavior*, *11*(1), e01936. (2) Nakai T, Koide-Majima N, Nishimoto S (2018) Encoding and decoding of music genre representations in the human brain. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.

Value of the Data

- Brain activity data with music genre labels to allow for supervised machine learning of music information.
- Neuroimaging and machine learning researchers can test multiple algorithms regarding the prediction and decoding performances of brain activity induced by various music stimuli.
- The large amount of data for each participant allows researchers to construct participant-wise modeling of brain activity.
- The test dataset with multiple repetitions of identical stimuli can be used to evaluate the signal to noise ratio of brain activity.
- Behavioral assessment of a genre recognition task allows for a comparison of categorical judgments based on behavior and brain activity.
- Formatting in the Brain Imaging Data Structure (BIDS) allows researchers to efficiently process the dataset.

1. Data Description

All raw brain data are available in OpenNeuro (https://openneuro.org/datasets/ds003720). The brain data are organized based on the BIDS [1], which makes it easier for researchers to preprocess and analyze brain data. The dataset includes anatomical magnetic resonance imaging (MRI) and functional MRI (fMRI) data. The behavioral data, presentation software script, and stimuli preprocessing scripts are available in the Open Science Framework (OSF) (https://osf.io/2zwkt/). The original music stimuli are available at (http://marsyas.info/downloads/datasets.html). These original stimuli are preprocessed for the experiments, as described in the **Materials and Methods** section.

Fig. 1 illustrates the experimental design and stimuli. Table 1 reports the age, sex, and number of fMRI runs and samples for each participant. Our experiment is based on a small-N

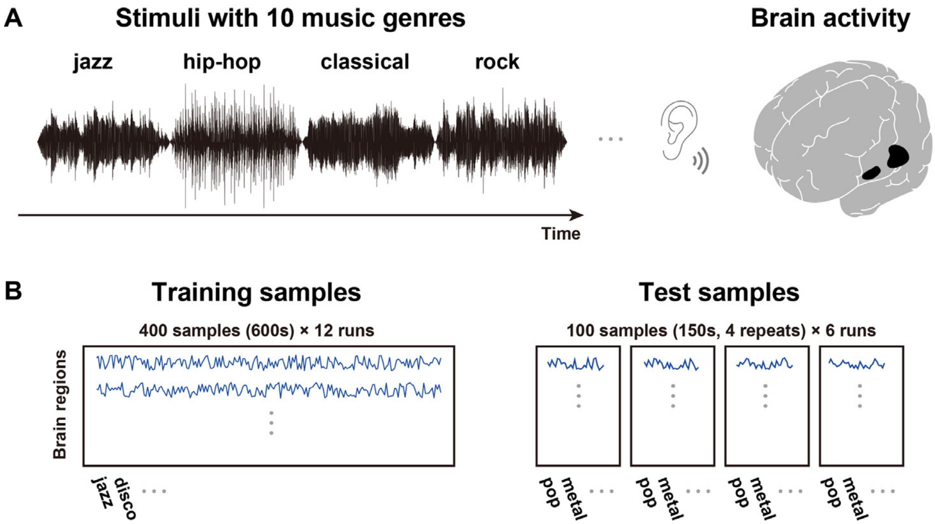

Fig. 1. Experimental design. (A) Subjects passively listened to 540 music pieces from 10 music genres. (B) The dataset is separated into training and test samples. Training samples were measured while the sequence of training music clips was presented once. Test samples were measured while the sequence of test music clips was presented four times. Modified from [3].

Table 1
Age, sex, and number of fMRI runs and samples for each participant.

	sub-001	sub-002	sub-003	sub-004	sub-005
Age	30	33	23	30	25
Sex	M	F	M	M	F
Music experience (years)	12	15	4	12	10
Total fMRI runs	18	18	18	18	18
Training samples (volumes)	4800	4800	4800	4800	4800
Test samples (volumes) (with four repetitions)	600	600	600	600	600

design (five subjects). We obtained 3 hours of neuroimaging data for each participant, allowing researchers to construct participant-wise modeling of brain activity. Although this dataset has been used in two previous studies [2,3], there remains much room to apply different acoustic models. The original music stimuli have been widely used in previous music information retrieval studies (see [4] for a review). Therefore, various hand-crafted models and neural network models can be tested using this dataset.

2. Experimental Design, Materials and Methods

2.1. Subjects

Five healthy subjects took part in the MRI and behavioral experiments (referred to as sub-001 to sub-005; aged between 23 and 33 years old; 3 males; music experience between 4 and 15 years). The subjects were all right-handed, with normal hearing, normal or corrected to normal vision, and no history of psychological or neurological disorders. No participant had absolute pitch. A questionnaire was used to assess the number of years that subjects trained on their primary instrument; this was used as the index of musical experience.

2.2. Stimuli and task

We obtained the original 1,000 music pieces from the GTZAN music genre dataset (http://marsyas.info/downloads/datasets.html) [5]. This dataset consists of 10 music genres (blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock). From the original music pieces, we randomly selected 54 music pieces (30 s, 22,050 Hz) from each music genre. A total of 540 music pieces were thus used in the current study. A 15 s music clip was randomly extracted from each music piece. We applied 2 s of fade-in and fade-out effects for each music clip. All music clips were normalized for their root mean square.

The fMRI experiment consisted of 12 training runs and 6 test runs (18 runs in total). Each run lasted 10 min and consisted of 40 music clips. At the beginning of each run, 15 s of dummy images were acquired to allow for the equilibration of the MRI signal; this should be omitted from further analysis. During the dummy scanning, the last music clip of the previous run was presented in the training runs, while the last clip of the same run was presented in the test runs. A total of 480 music clips were used for the training runs, and the left-out 60 music clips were used for the test runs. In each test run, a set of 10 music clips (150 s in total) was presented four times in the same order. There was no repetition in the training runs. During scanning, subjects were instructed to fixate on a crosshair at the center of the screen and to listen to the music stimuli through MRI-compatible insert earphones (Model S14, Sensimetrics). This type of earphone has been widely used in previous MRI studies to attenuate scanner noise [7]. After each 10 min run, the subjects were asked to describe their physical condition. A 1–2 min break was allowed if they felt fatigued or sleepy. At the end of the experiment of each day, the subjects were asked whether they fell asleep during the fMRI experiment. According to their self-reports, no subjects slept during the experiments. The MRI experiment was performed over the course of three days, with six fMRI runs performed each day. Presentation software (Neurobehavioral Systems, Albany, CA, USA) was used to control the stimulus presentation and the collection of behavioral data. The reference anatomical scans were also obtained on the first day.

2.3. MRI data acquisition

We used a 3.0 T MRI scanner (TIM Trio; Siemens, Erlangen, Germany). A 32-channel head coil was used. All experiments were performed at the Center for Information and Neural Networks (CiNet), National Institute of Information and Communications Technology (NICT), Osaka, Japan. For functional scanning, we scanned 68 interleaved axial slices with a thickness of 2.0 mm without a gap using a T2*-weighted gradient echo multi-band echo-planar imaging (MB-EPI) sequence [6] (repetition time (TR) = 1500 ms, echo time (TE) = 30 ms, flip angle (FA) = 62°, field of view (FOV) = 192 × 192 mm^2, voxel size = 2 × 2 × 2 mm^3, MB factor = 4). We obtained 410 volumes (615 s) for each run. We also used a magnetization prepared rapid acquisition gradient echo sequence (MPRAGE) to acquire high-resolution T1-weighted images of the whole brain (TR = 2530 ms, TE = 3.26 ms, FA = 9°, FOV = 256 × 256 mm^2, voxel size = 1 × 1 × 1 mm^3).

2.4. Behavioral experiment

To compare brain activation patterns and behavioral performance, we asked the subjects to participate in additional behavioral experiments. The behavioral experiments were conducted in a soundproof room. Subjects were first instructed to listen to three original music clips (30 s) for each music genre as a reference. These music clips were selected randomly from the 460 music pieces which were not used in the MRI experiment. The subjects were provided the information on the correct music genres in this training session. The subjects then listened to the 60 music clips which had been used as the test dataset in the MRI experiment. The subjects selected the music genre of the target music clip by filling in 1 of 10 cells on the answer sheet displayed on a laptop computer screen. The music clips were presented through closed headphones. Music clips were presented in the same order as in the MRI experiments. Subjects were allowed to listen to each music clip only once.

Ethics Statement

We obtained informed consent from all subjects before their participation. The current experiment was approved by the ethics and safety committees of the NICT in Osaka and Tokyo, Japan (protocol number: 194104).

CRediT Author Statement

Tomoya Nakai: Conceptualization, Methodology, Software, Investigation, Software, Visualization, Data curation, Formal analysis, Writing-Original draft; **Naoko Koide-Majima:** Conceptualization, Investigation, Visualization, Writing-Reviewing, and Editing; **Shinji Nishimoto:** Writing-Reviewing and Editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially supported by MEXT/JSPS KAKENHI (grant number JP20K07718, JP20H05023 in #4903 (Evolinguistics), JP15H05311) as well as JST CREST JPMJCR18A5 and ERATO JPMJER1801. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] K.J. Gorgolewski, T. Auer, V.D. Calhoun, R.C. Craddock, S. Das, E.P. Duff, G. Flandin, S.S. Ghosh, T. Glatard, Y.O. Halchenko, D.A. Handwerker, M. Hanke, D. Keator, X. Li, Z. Michael, C. Maumet, B.N. Nichols, T.E. Nichols, J. Pellman, J.B. Poline, A. Rokem, G. Schaefer, V. Sochat, W. Triplett, J.A. Turner, G. Varoquaux, R.A. Poldrack, The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments, Sci. Data 3 (2016) 160044, doi:10.1038/sdata.2016.44.

[2] T. Nakai, N. Koide-Majima, S. Nishimoto, Encoding and decoding of music-genre representations in the human brain, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 2018.

[3] T. Nakai, N. Koide-Majima, S. Nishimoto, Correspondence of categorical and feature-based representations of music in the human brain, Brain Behav. 11 (2021) e01936.

[4] B.L. Sturm (2013) The GTZAN dataset: Its contents, its faults, their effects on evaluation, and its future use arXiv preprint arXiv:1306.1461.

[5] G. Tzanetakis, P. Cook, Musical genre classification of audio signals, IEEE Trans. Speech Aud. Process 10 (2002) 293–302.

[6] S. Moeller, E. Yacoub, C.A. Olman, E. Auerbach, J. Strupp, N. Harel, K. Uğurbil, Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain FMRI, Magn. Reson. Med. 63 (2010) 1144–1153.

[7] S. Norman-Haignere, N.G. Kanwisher, J.H. McDermott, Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition, Neuron 88 (2015) 1281–1296, doi:10.1016/j.neuron.2015.11.035.