



ELSEVIER

Contents lists available at ScienceDirect

Data in brief

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



Data Article

A dataset for evaluating one-shot categorization of novel object classes



Yaniv Morgenstern^{*}, Filipp Schmidt, Roland W. Fleming

Department of Experimental Psychology, Justus-Liebig University Giessen, Giessen, 35394, Germany

ARTICLE INFO

Article history:

Received 9 November 2019

Received in revised form 2 February 2020

Accepted 10 February 2020

Available online 21 February 2020

Keywords:

One-shot learning

Categorization

Generalization

Abstraction

Machine vision

Objects

Visual perception

Shape

ABSTRACT

With the advent of deep convolutional neural networks, machines now rival humans in terms of object categorization. The neural networks solve categorization with a hierarchical organization that shares a striking resemblance to their biological counterpart, leading to their status as a standard model of object recognition in biological vision. Despite training on thousands of images of object categories, however, machine-learning networks are poorer generalizers, often fooled by adversarial images with very simple image manipulations that humans easily distinguish as a false image. Humans, on the other hand, can generalize object classes from very few samples. Here we provide a dataset of novel object classifications in humans. We gathered thousands of crowd-sourced human responses to novel objects embedded either with 1 or 16 context sample(s). Human decisions and stimuli together have the potential to be re-used (1) as a tool to better understand the nature of the gap in category learning from few samples between human and machine, and (2) as a benchmark of generalization across machine learning networks.

© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

DOI of original article: <https://doi.org/10.1016/j.visres.2019.09.005>.

^{*} Corresponding author. Department of Experimental Psychology, Justus-Liebig University Giessen, Otto-Behaghel-Str.10F, Giessen, 35394, Germany.

E-mail address: Yaniv.Morgenstern@psychol.uni-giessen.de (Y. Morgenstern).

<https://doi.org/10.1016/j.dib.2020.105302>

2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Specifications Table

Subject	Experimental and Cognitive Psychology
Specific subject area	Vision and Perception
Type of data	MATLAB code (that analyses raw data) Raw Data Visual Stimuli (as *.png, and as generated with MATLAB) Text File, describing the shared contents PNG files that show the output of the MATLAB analysis routines
How data were acquired	Clickworker (crowd-sourcing platform)
Data format	Raw data (*.mat file) and MATLAB code to analyse raw data
Parameters for data collection	Independent factors (see also Fig. 1) <i>cont_num</i> - The number of context samples (1 or 16) <i>var_num</i> - The variability of the context samples (low or high) <i>obj_num</i> - novel objects (20 classes) <i>distbin_num</i> - the distance of the test to the underlying base shape (1 of 25 bins) <i>shape_num</i> - shape identification number in each <i>distbin_num</i> (1–10). dependent factors Human yes/no responses on whether a central test belongs to the same class as the sample shape(s)
Description of data collection	Data were collected with an online crowd-sourcing platform (clickworker). In the main experiment, 500 observers responded to 22 trials indicating whether a central test was in the same class as the samples, which consisted of 1 or 16 sample shapes that varied in their similarity to the test. Fig. 1 shows 4 example trials. For each participant, the first 20 trials showed a different novel shape. The last two trials were catch trials with the same central test and surrounding sample shapes.
Data source location	Giessen, Hessen
Data accessibility	Repository name: Zenodo Data identification number: 10.5281/zenodo.3628659 Direct URL to data: http://doi.org/10.5281/zenodo.3628659
Related research article	Morgenstern, Y., Schmidt, F., & Fleming, R. W. (2019). One-shot categorization of novel object classes in humans. <i>Vision Research</i> , 165, 98–108. DOI: https://doi.org/10.1016/j.visres.2019.09.005

Value of the Data

- From just a single example, we can derive quite precise intuitions about what other class members look like. This stands in stark contrast to machine learning algorithms, which typically require tens or even hundreds of thousands of examples to learn a new category. One of the most important open questions in our field is: How do humans achieve this? The stimuli and data provided in this paper can be used to test machine learning generalization as compared to human and also can be used as a test bed for various kinds of category learning models.
- These data will benefit cognitive and machine learning scientists interested in testing how their category learning theories and algorithms transfer to the human perception of novel object classes from few samples.
- Specifically, these stimuli can be analyzed with a model according to an experimenter's preference (e.g., by calculated similarity between the test and sample(s) in terms of image computable features) to produce responses that can be compared to human responses. In this way, the data can be used to test an experimenter's theoretical ideas on how humans learn from few samples.
- These data can also be used as stimuli to examine category learning in the human brain (e.g., using fMRI to investigate brain activity changes as a function of test similarity to the samples)

1. Data

As the data includes stimuli and corresponding human responses, they allow an experimenter to run their model on the stimuli, and compare the result to human performance. For convenience, the shared MATLAB files, listed below, show how an experimenter can recover the shapes used in the stimulus frame of single experimental trial, and the corresponding human response to that trial.

1.1. Stimuli

The MATLAB script *oneSHOTdata_demo.m* creates two figure windows. One figure draws the stimulus frame for a particular trial (Fig. 1). Another figure draws the test (in blue) and sample(s) (in black) and displays the average distance (in terms of skeletal parameters) of the test shape with the sample(s), and the human response (in terms of whether the test is in the same class as the sample; 0 = no, 1 = yes).

The initial parameters for this script are the object identification number (*obj_num*; 1–20), the number of samples (*cont_num*; 1 or 16), the distance between the base shape used to create the stimuli and the test shape (*distbin_num*; 1–25), and the shape identification number for each distance bin (*shape_num*; 1–10).

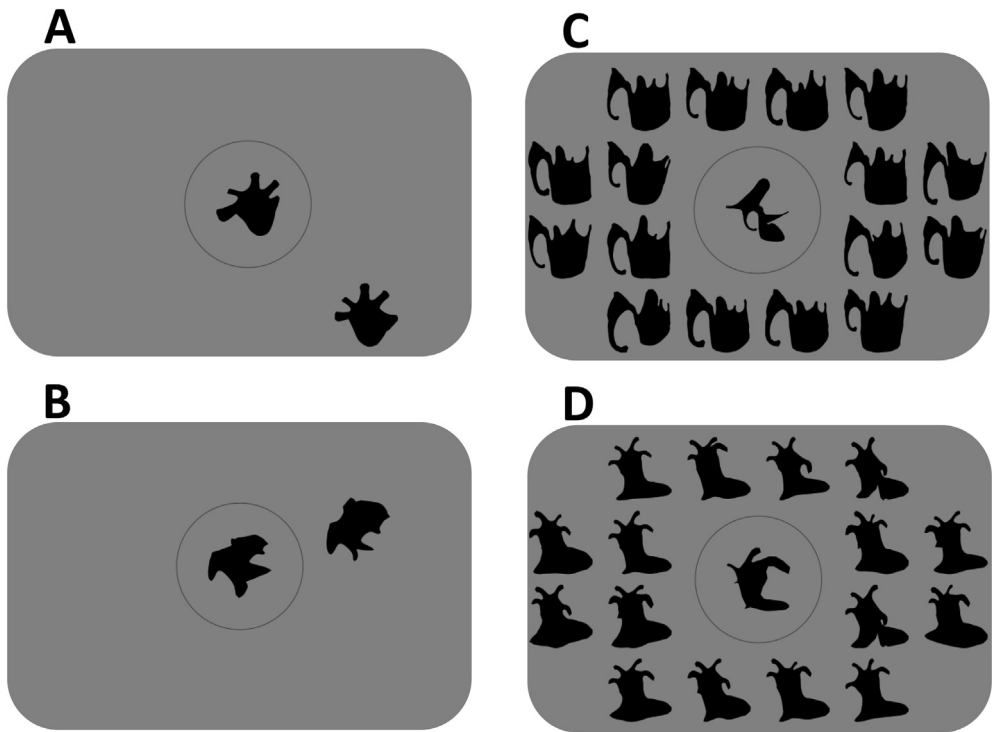


Fig. 1. Task and Example Stimuli. Crowdsourced observers judged whether the test (embedded in circle) was in the same categorical class as the surrounding sample(s). In 20 trials, each observer (500 total) judged unique combinations of test and sample(s) to 20 novel object classes (only 4 shown). On each trial, the stimulus frame showed the test embedded in 1 (A, B) or 16 (C, D) surrounding samples in object classes that tended to vary more (C, B) or less (A, D). The test itself varied in its appearance from the samples in 1 of 25 bins based on its skeletal similarity to an original base shape (see *Stimulus Frame in Experimental Design, Materials, and Methods*). The stimulus frames and human responses are available through Zenodo. By varying data collection parameters listed in the Specifications Table (*obj_num*, *cont_num*, *distbin_num*, and *shape_num*), the MATLAB code *oneSHOTdata_DEMO.m* shows how to generate a stimulus frame and recover the corresponding human response from the raw data for an individual trial number. As examples, stimulus frame (A) was rendered with object number 2 (*obj_num* = 2, which had surrounding samples that tended to appear more similar, lower variability), with only 1 surrounding sample (*cont_num* = 1), with a test that had a smaller skeletal parameter distance from the base shape (*distbin_num* = 8), and a shape 3 of 10 for this condition (*shape_num* = 3) while stimulus frame (B) was rendered with object number 1 (*obj_num* = 1, with higher variability across samples), with 16 surrounding samples (*cont_num* = 16), with a test that had a large skeletal parameter distance from the base shapes (*distbin_num* = 25), and was shape 2 out of 10 for this condition (*shape_num* = 2). In (C) *obj_num* = 1, *cont_num* = 16, *distbin_num* = 25, and *shape_num* = 1. In (D) *obj_num* = 5, *cont_num* = 1, *distbin_num* = 12, and *shape_num* = 6.

This script can be used to extract shapes (shapes), human responses (*hresp*; 0 = no, 1 = yes), the average distance of the test to the sample(s) (*adist*), and the sample variability condition (*var_num*; 1 = low, 2 = high).

1.2. Human responses

Also, the human responses can be extracted from *oneSHOTdata_DEMO.m*. Other files reproduce the data figures in Morgenstern, Schmidt, and Fleming [1], specifically Figs. 3a (script: *Fig3a.m*), 4a (script: *Fig4a.m*), 4b (script: *Fig4b.m*), 4c (script: *Fig4c.m*), 5 (script: *Fig5.m*), and 7 (script: *Fig7.m*).

All of these scripts load a MATLAB structure ('*data \ 1shot_data.mat*') with the raw data that has the relevant information for each trial, including:

- the responses (*data.resp*),
- shape dissimilarity (*data.actual_dist_center*) – defined as the average distance of test to the samples,
- the object number (*data.obj_id*),
- the number of samples in the surround (*data.contxt_id*),
- the participant id (*data.subj*).

In addition, to create Fig. 7 from Ref. [1], script *Fig7.m* loads data for a control experiment that measures correlation across human observers ('*data/hum_consistency_data.mat*' in *estimate_hh_con.m*).

Figs. 4 and 5 from Ref. [1] plot the data averaged across objects. However, Fig. 3a plots data for individual objects as probability YES as a function of shape dissimilarity in the one (light colour) and many (dark colour) sample conditions. Varying the *obj_num* variable from 1 to 20 in script *Fig3a.m* will produce Fig. S5 in Ref. [1].

2. Experimental design, materials, and methods

2.1. Stimulus generation

Twenty novel base shapes were created as described in Ref. [1]. Then, we generated novel samples by transforming the base shape's skeletal representation [2] to synthesize new shapes with sub-parts (limbs) that varied in terms of their spatial position, length, width, and orientation.

2.2. Stimulus frame

The experimental stimulus frame for each base shape includes a test embedded in 1 or 16 sample(s) (Fig. 1; see *oneSHOTdata_DEMO.m* for an example). The test varied in terms of the Euclidean distance of its underlying skeletal parameters to that of the base shape. The test distance ranged from near to the base shape (*distbin_num* = 1; in *oneSHOTdata_DEMO.m*) to far (*distbin_num* = 25) in 25 equal sized bins. For each bin, there were 10 stimulus frames (that varied in terms of the shapes for the test and samples).

2.3. Veridical distance

The distribution of sample shapes was selected to be near the base shape in terms of their skeletal parameters (see Supplemental Methods S1 in Ref. [1]) to produce sample shapes that appear similar to the base shape. Thus, for any given trial, the surrounding samples were on average near the parameters for the base shape rather than equal to the base shape. Thus, for our analysis, we do not use the distance in skeletal parameters of the base shape to the test shape. Rather, to more precisely portray the distance between test and samples, we use the average distance between the skeletal parameters of the test and the surrounding samples (scripts: *Fig3a.m*, *Fig4a.m*, *Fig4b.m*, *Fig4c.m*, *Fig5.m*, *Fig7.m*).

2.4. Sample variability

Base shapes with more sub-parts (limbs) produced novel shape samples that appeared to show greater class variability (see Figs. 2, S3, and S4 in Ref. [1]). Thus, base shapes with 6 or more limbs were grouped as highly variability samples and base shapes with 5 or less limbs were grouped as low variability samples. MATLAB script *Fig4a.m* shows how to determine whether an object is in the low or high variability group. The skeletal structure for each base shape is embedded in *stim \ novobj.s.mat* and can be visualized using MATLAB code provided by Ref. [2].

Acknowledgments

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 222641018-SFB/TRR 135 TP C1 and the ERC Consolidator award “SHAPE” (ERC-CoG-2015-682859). We thank Brendan Nicholls and Hendrik Will for their help setting up the experiments and running initial pilot studies.

Conflict of 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.dib.2020.105302>.

References

- [1] Y. Morgenstern, F. Schmidt, R.W. Fleming, One-shot categorization of novel object classes in humans, *Vis. Res.* 165 (2019) 98–108.
- [2] J. Feldman, M. Singh, Bayesian estimation of the shape skeleton, *Proc. Natl. Acad. Sci. Unit. States Am.* 103 47 (2006) 18014–18019.