**MDPE dataset description**

The Multimodal Deception Dataset with Personality and Emotional Characteristics (MDPE) consists of three parts: personality characteristics, emotional characteristics, and deception characteristics.

﻿

**Personality Characteristics**: Subjects were required to fill out a Big Five personality questionnaire, which consists of 60 questions. Each question was marked with a score indicating whether the descriptions match their own, with 1 indicating strong disagreement and 5 indicating strong agreement. Details can be found in the “Big Five Personality Inventory Second Edition (BFI-2).docx”. The “personality.csv” file contains the personality traits of the subjects.

﻿

**Emotional Characteristics**: Subject were asked to watch 16 emotion induction videos (sadness, happiness, relaxation, surprise, fear, disgust, anger, neutral, two Inducing videos for each emotion in random order). Subjects were required to describe their feelings after watching and fill out an emotion scale in the form of self-report (quantifying 8 emotions, with 1-5 points for each emotion representing intensity, 1 indicating no such emotion, and 5 indicating the strongest emotion).

For audio and text modalities, due to some sensitive factors, **we only released features instead of the raw audio and transcribed text**. The features here only includes the emotional part of the theme description. (The subject did not speak while watching the induction video)

﻿

**Deception Characteristics**: The subjects were asked 24 questions, of which 9 were pre selected and had time to prepare (deception). During the interview process, when asked these 9 questions, the subject must lie, and when asked the remaining 15 questions, they must tell the truth. We provided subjects with effective monetary incentives to generate effective deceptive behavior.

During the interview process, the interviewer provided their judgment of truthful or deceptive answers to each question, record the trust level the interviewer thinks of each answer. The trust level was divided into 1-5 points, where 1 represents definitely true and 5 represents definitely false ("Interviewer evaluation" in "deception/lables/xxx.csv").

After the interview, the subject were asked to rate the answer they just lied to. The same score is 1-5, where 1 represents that I have definitely deceived successfully and 5 represents that I have definitely not deceived successfully ("Self evaluation of interviewee" in "deception/lables/xxx.csv", and 0 indicates that the question is not a deceptive problem).

During the interview process, the order of asking questions is random("Question id" in "deception/lables/xxx.csv"). The specific question numbers are as follows:

Interview Questions

1. What color do you like the most? Why?

2. Where is your hometown? Please briefly introduce it.

3. Do you have any hobbies?

4. Have you traveled in the past year?

5. How do you like Beijing?

6. What is your happiest experience?

7. What is your favorite food?

8. What is your personality like?

9. What is your biggest weakness?

10. What is your greatest strength?

11. What do you usually do to relax?

12. Which exercise or sport do you like?

13. Briefly introduce your family members.

14. Who is the person you have the greatest influence on you?

15. Do you have any special places or tourist destinations you want to go to?

16. Who is your favorite celebrity or great person?

17. What is your opinion on the words "neijuan" and "tangping"?

18. What is your favorite literary and artistic work?

19. Have you ever received any rewards or honors in school or at work?

20. What was your most unforgettable experience in the past year?

21. Have you participated in any major event?

22. Have you ever cheated in school or work?

23. Have you concealed a fact to your family or friends in the past year?

24. Have you ever lied to avoid responsibility?

The naming of deception features is “ID-X-Y”，where X is the order in which questions are asked and Y is the specific question number.

For more details about MDPE, please refer to:

[1] *Cai C, Liang S, Liu X, et al. MDPE: A Multimodal Deception Dataset with Personality and Emotional Characteristics[J]. arXiv preprint arXiv:2407.12274, 2024.*

Baseline Code: <https://github.com/cai-cong/MDPE>

Dataset download link:<https://huggingface.co/datasets/MDPEdataset/MDPE_Dataset>

EULA link: <https://drive.google.com/file/d/1A1F8szMOTf9-rK8DYD23GruBArtnYdLl/view?usp=sharing>

The data are passcode protected. Please download and send the signed EULA to mdpe.contact@gmail.com for access request.