Multimodal Algorithmic Reasoning - SMART-101 Challenge

Held in conjunction with Multimodal Algorithmic Reasoning Workshop at CVPR 2024
https://marworkshop.github.io/cvpr24/index.html

CALL FOR PARTICIPATION

In the last couple of years, we have seen dramatic improvements in the reasoning abilities of multimodal and large language models. In this SMART-101 CVPR 2024 challenge, we attempt to understand these abilities of large deep models through solving visuo-linguistic puzzles that need basic mathematical and algorithmic skills; these skills even young children seem to possess and can solve the puzzles without much difficulty. A thorough empirical analysis of such abilities of multimodal LLMs is the basic premise of our CVPR 2023 paper titled Are Deep Neural Networks SMARTer than Second Graders? This paper introduces the Simple Multimodal Algorithmic Reasoning Task (SMART) and the SMART-101 dataset. Building upon the efforts in the paper, this SMART-101 CVPR-2024 challenge is an attempt at bringing research interest into this important topic to understand where we stand in the race towards achieving true Artificial General Intelligence (AGI). Specifically, the goals of this competition are three-fold, towards understanding:
(i) how well do state-of-the-art multimodal LLMs abstract data, attend to key details, and generalize their knowledge to solve new problems?
(ii) how fluid are they in acquiring new skills? and
(iii) how effective are they in the use of language for visual reasoning?

Through the state-of-the-art AI models submitted by the participants of this challenge, we hope to learn and understand where we stand in real AGI abilities, and more importantly, clearly answer if the current AI is at least better than second graders in mathematical/algorithmic abilities.

The SMART challenge involves solving visuo-linguistic puzzles designed specifically for children in the 6–8 age group. The puzzles are taken from the Math Kangaroo Olympiad -- a popular international children's Olympiad that uses a multiple choice answer selection format. Most of the puzzles have an image and a text question, and five answer options of which only one option is the correct answer to the puzzle. Participant submissions to the challenge will be evaluated against a private test set. The solution to each puzzle needs a mix of various basic mathematical and algorithmic reasoning skills, involving

basic arithmetic, algebra, spatial reasoning, logical reasoning, measuring, path tracing, pattern matching, and counting.

IMPORTANT DATES

* SMART-101 Challenge Track
        Challenge open: March 28, 2024
        Submission deadline: ***June 8, 2024 7:59:59 PM EST (GMT - 4:00)***
        Arxiv paper submission deadline: June 10, 2024 (AOE time)
        Public winner announcement: June 17, 2024

INSTRUCTIONS FOR PARTICIPATING IN THE SMART-101 CHALLENGE

---

* The challenge is hosted on Eval.AI. Please read the instructions at the following link for the submission guidelines: https://eval.ai/web/challenges/challenge-page/2247/overview

* The challenge participants are required to make arXiv submissions detailing their approach as well as make their implementation publicly available on Github to be considered for the prizes. Note that the participant's arXiv submissions will not be part of the workshop proceedings.

* Winners of the challenge are determined both by the performance on the leaderboard over a private test set as well as the quality of the proposed method (as detailed in their arXiv submission and reviewed by a panel). Please see the details on the challenge website.

* Prizes will be awarded on the day of the workshop.

WORKSHOP ORGANIZERS

Anoop Cherian, Mitsubishi Electric Research Laboratories
Suhas Lohit, Mitsubishi Electric Research Laboratories
Honglu Zhou, Salesforce Research
Moitreya Chatterjee, Mitsubishi Electric Research Laboratories
Kuan-Chuan Peng, Mitsubishi Electric Research Laboratories
Kevin A. Smith, Massachusetts Institute of Technology
Tim K. Marks, Mitsubishi Electric Research Laboratories
Joanna Matthiesen, Math Kangaroo USA
Joshua B. Tenenbaum, Massachusetts Institute of Technology

CONTACT
Email: smart101@googlegroups.com

SMART-101 Challenge: https://eval.ai/web/challenges/challenge-page/2247/overview
Website: https://marworkshop.github.io/cvpr24/