Advancing AI that is more aware of our emotional and cultural context

December 14 2022

ArtELingo, a multilingual dataset and benchmark of WikiArt with captions & emotions. Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2211.10780

King Abdullah University of Science and Technology (KAUST) has shown for the first time with a large-scale dataset how emotions may vary in response to visual stimuli over several languages and cultures.

ArtELingo is a multilingual dataset with emotional explanations constructed from exposure to 80,000 visual stimuli (visual artworks) in multiple languages. Its current version, accepted to the dataset, includes more than 420,000 of these annotations for each of the English, Chinese, and Arabic languages. A small version was also collected in Spanish on over 1000 artworks from Latin America and Latin Europe to explore how two different cultures speaking the same language may vary in terms of the constructed emotions.

The ArtELingo paper was presented at The 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP).

ArtELingo is a step towards creating culturally diverse datasets that represent both western and non-western cultures well. This will particularly help those looking for data when studying cultural and cross-cultural psychology. Overall, this research helps advance building more human-compatible AI aware of our emotional and cultural beings.

This project has been developed by Youssef Mohamed, Mohamed Abdelfattah, Shyma Alhuwaider, Feifan Li, and Mohamed Elhoseiny (PI) from KAUST, and collaborators from the University of Notre Dame (Xiangliang Zhang) and Kenneth Ward Church (NorthEastern university).

The authors wish to thank Baidu, Beijing time for their support in collecting the Chinese version of the dataset, and tens of universities from Egypt (mainly) and Saudi Arabia for collecting the Arabic version.

The work is published on the *arXiv* preprint server.

Code repository: github.com/Vision-CAIR/artelingo

Provided by King Abdullah University of Science and Technology