**Why Does a Visual Question Have Different Answers?**

Nilavra Bhattacharya 1, Qing Li 2, Danna Gurari 1

1 The University of Texas at Austin, 2 University of California, Los Angeles

Motivation: Different People Can Offer Different Answers to a Visual Question
- Approximately a 50/50 split for when a single versus multiple answers are observed for ~500,000 visual questions in three datasets [1]

Goal: Understand Why Answers Differ
- Benefits include empowering system designers and users with knowledge in how to interpret, prevent, and resolve answer differences
- More generally, support developing algorithmic frameworks that can account for the diversity of perspectives in a crowd

References

**Answer-Differences Dataset**

**Taxonomy of nine reasons for why answers can differ:**

- Mentality: Is there a preference for a certain color among people?
- Option: The possible set of answers to the question.
- Ambiguity: The image is ambiguous.
- VQA answerability: The question is unanswerable.
- Random: The question is random.
- Relevance: The question is irrelevant.
- Ambiguity, synonymous answers, & varying levels of answer granularity
- VQA answerability

**Frequency of each reason leading to answer differences:**

- VizWiz [2] (29,291 VQs)
- VQA_2.0 [3] (15,034 VQs)

- Most common reason: ambiguity, synonymous answers, & varying levels of answer granularity
- Least common reason: spam (~1% of VQs)

**Novel Task and Algorithm: Given a Visual Question, Predict the Reason(s) Why Answers from a Crowd Will Differ**

**Algorithm architecture:** predicts with image, question, & predicted answers

**Evaluation:** mean average precision for all reasons

**Model**

**Overall**

**Model**

**Overall**

Random 30.15
VQA: when VQ is irrelevant, then LQ, IV, & AMB are reasons
Relevance [4] 31.71
QI-Relevance [4] 32.23
VQ answerability: when VQ is unanswerable, then LQ, IV, & AMB are reasons
Random 31.88
Q 43.47
Q + I 43.16
Q + I, FT 44.55
Q + I, GT 44.46
Q + I, FT 44.09

**DATASET:** https://vizwiz.org/

**DATASET:** https://vizwiz.org/