We Are Humor Beings: Understanding and Predicting Visual Humor

Arjun Chandrasekaran¹, Ashwin K. Vijayakumar¹, Stanislav Antol¹, Mohit Bansal², Dhruv Batra¹, C. Lawrence Zitnick³, Devi Parikh¹

¹Virginia Tech, ²Toyota Technological Institute, ³Facebook AI Research

Problem Statement

**Task 1: Predict the funniness score of a scene**

- Predict funny pictures to upload to social media.
- Generate funny scenes for entertainment.
- Personalized recommendation catering to specific preferences of humor type.

**Task 2: Alter funniness of a scene**

- Replace object / don't replace object
- Support Vector Regression
- MLP (3 hidden layers)

Abstract Visual Humor (AVH) dataset

Enables scene-level understanding of humor.

**AMT task:**

Create scenes that are funny yet plausible.

**Funniness score = Avg. rating**

Inter-human agreement on funniness score

Incongruity theory: Humor is induced “when a stimulus violates an expectation”

80% of the scenes displayed unusual humor techniques present in the scene.

Investigated the humor techniques present in a subset of 200 scenes.

**Task 1:**

Predict the funniness score

Step 1: Find objects to be replaced

Support Vector Regression

MLP (3 hidden layers)

Step 2: Replace an object with another object

Quantitative Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SVR</td>
<td>0.24</td>
<td>24.53%</td>
</tr>
<tr>
<td>Average Prediction</td>
<td>0.315</td>
<td>7.69%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Class</th>
<th>Class Avg. Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do not replace</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Anomaly detection</td>
<td>83.91%</td>
<td></td>
</tr>
<tr>
<td>Our model</td>
<td>74.45%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 1: Find objects to replace</th>
<th>Step 2: Replace an object with another object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replace object / don't replace object</td>
<td>Replace object / don't replace object</td>
</tr>
<tr>
<td>Support Vector Regression</td>
<td>MLP (3 hidden layers)</td>
</tr>
</tbody>
</table>

Human evaluation

**Type of evaluation:**

-1.05 (Funniness score decreases)

95% less funny

**Relative**

*Compared whose scene is funnier: Human scene vs model’s scene

Discussion

**Humor vs anomaly**

- Supervised humor detection outperforms unsupervised anomaly detection.
- Thus, an oddity (incongruity) may result in humor but not all oddities are funny.
- Question: What are the necessary and sufficient conditions for humor?

**Perception of real vs abstract scenes**

- "Suspension of reality" in abstract scenes makes the "benign" violations of expectations more funny.

**Role of captions in humor**

- "What impact do captions have in the perceived funniness of a scene?"

**Mental model” in visual humor**

- "The punch” at the end of a joke breaks our expectations that are formed during the initial “setup”.
- Question: What is our mental model that is violated in the case of visual humor? How is it formed?

Data-efficient method that is expressive and does not require large amounts of human-annotated data.

"... raccoons have a drinking problem."