

<h1>SugarCrepe</h1>	<p>SugarCrepe is a benchmark for faithful vision-language compositionality evaluation. Given an image, a model is required to select the <i>positive</i> text that correctly describes the image, against another <i>hard negative</i> text distractor that differs from the positive text only by small compositional changes.</p>
<p><b>DATASET LINK</b></p> <p><b>Dataset:</b>  <a href="https://github.com/RAIVNLab/sugar-crepe">github.com/RAIVNLab/sugar-crepe</a></p>	<p><b>DATA CARD AUTHOR(S)</b></p> <p><b>Cheng-Yu Hsieh*, University of Washington</b></p> <p><b>Jieyu Zhang*, University of Washington</b></p> <p><b>Zixian Ma, University of Washington</b></p> <p><b>Aniruddha Kembhavi, Allen Institute for Artificial Intelligence</b></p> <p><b>Ranjay Krishna, University of Washington</b></p>

Dataset Overview		
DATA SUBJECT(S)	DATASET SNAPSHOT	CONTENT DESCRIPTION
Sensitive Data about people  <b>Non-Sensitive Data about people</b>  <b>Data about natural phenomena</b>  <b>Data about places and objects</b>  <b>Synthetically generated data</b>  Data about systems or products and their behaviors  Unknown  Others (Please specify)	Size of Dataset	1.7 MB
	Number of Instances	7512
	Number of Fields	3
	Labeled Classes	N/A
	Number of Labels	N/A
	Average Labels Per Instance	N/A
	Algorithmic Labels	N/A
	Human Labels	N/A
	Other Characteristics	N/A
<b>Sensitivity of Data</b>		
SENSITIVITY TYPE(S)	FIELD(S) WITH SENSITIVE DATA	SECURITY AND PRIVACY HANDLING

User Content User Metadata User Activity Data Identifiable Data S/PII Business Data Employee Data Pseudonymous Data Anonymous Data Health Data Children's Data <b>None</b> Others (Please specify)	N/A	N/A
RISK TYPE(S)	SUPPLEMENTAL LINK(S)	RISK(S) AND MITIGATION(S)
Direct Risk Indirect Risk Residual Risk <b>No Known Risks</b> Others (Please Specify)	N/A	N/A
<b>Dataset Version and Maintenance</b>		
MAINTENANCE STATUS	VERSION DETAILS	MAINTENANCE PLAN
Regularly Updated (New versions of the dataset have been or will continue to be made available.)	<b>Current Version:</b> 1.0 <b>Last Updated:</b> 06/2023 <b>Release Date:</b> 06/2023	We are committed and have resources to maintain the dataset. <b>Versioning:</b> N/A <b>Updates:</b> Dataset may be updated to address technical issues. <b>Errors:</b> We actively monitor issues at <a href="https://github.com/RAIVNLab/sugar-crepe">github.com/RAIVNLab/sugar-crepe</a>

<p><b>Actively Maintained</b> (No new versions will be made available, but this dataset will be actively maintained, including but not limited to updates to the data.)</p> <p><b>Limited Maintenance</b> (The data will not be updated, but any technical issues will be addressed.)</p> <p><b>Deprecated</b> (This dataset is obsolete or is no longer being maintained.)</p>		<p><b>Feedback:</b> We actively monitor issues at <a href="https://github.com/RAIVNLab/sugar-crepe">github.com/RAIVNLab/sugar-crepe</a></p>
	NEXT PLANNED UPDATE(S)	EXPECTED CHANGE(S)
	N/A	N/A

Example of Data Points														
PRIMARY DATA MODALITY	SAMPLING OF DATA POINTS	DATA FIELDS												
Image Data Text Data Tabular Data Audio Data Video Data Time Series Graph Data Geospatial Data <b>Multimodal (Image and Text Data)</b> Unknown Others (Please specify)	[Typical Data Point Link]  <a href="https://github.com/RAIVNLab/sugar-crepe/tree/main/data">github.com/RAIVNLab/sugar-crepe/tree/main/data</a>	<table> <tr> <th>Field Name</th><th>Field Value</th><th>Description</th></tr> <tr> <td>filename</td><td>string</td><td>The id to an image</td></tr> <tr> <td>caption</td><td>string</td><td>Positive text correctly describing the image</td></tr> <tr> <td>negative_caption</td><td>string</td><td>Hard negative text incorrectly describing the image</td></tr> </table>	Field Name	Field Value	Description	filename	string	The id to an image	caption	string	Positive text correctly describing the image	negative_caption	string	Hard negative text incorrectly describing the image
Field Name	Field Value	Description												
filename	string	The id to an image												
caption	string	Positive text correctly describing the image												
negative_caption	string	Hard negative text incorrectly describing the image												
TYPICAL DATA POINT		ATYPICAL DATA POINT												
<pre>{   "filename": "000000289393.jpg",   "caption": "Several toy animals - a bull, giraffe, deer and parakeet.",   "negative_caption": "Several toy animals - a bull, giraffe, snake and parakeet." }</pre>		N/A												

# Motivations & Intentions

## Motivations

PURPOSE(S)	DOMAIN(S) OF APPLICATION	MOTIVATING FACTOR(S)
Monitoring Research Production Others (Please Specify)	Machine Learning, Computer Vision, Natural Language Processing, Vision-Language Models, Compositional Understanding	<ul style="list-style-type: none"><li>Faithfully evaluating compositionality of vision-language models.</li></ul>

## Intended Use

DATASET USE(S)	SUITABLE USE CASE(S)	UNSUITABLE USE CASE(S)
Safe for production use Safe for research use Conditional use-some unsafe applications Only approved use Others (Please specify)	<ul style="list-style-type: none"><li>Evaluate the compositionality of vision-language models.</li></ul>	N/A

LICENSING	RESEARCH AND PROBLEM SPACE(S)	CITATION GUIDELINES
We license our work using <a href="#">MIT License</a> . All the source data we use is publicly released by prior work.	Evaluate the compositionality of vision-language models.	<b>BiBTeX:</b> <pre>``` @article{hsieh2023sugarcrape,   title={SugarCrepe: Fixing Hackable Benchmarks for Vision-Language Compositionality},   author={Hsieh, Cheng-Yu and Zhang, Jieyu and Ma, Zixian and Kembhavi, Aniruddha and Krishna, Ranjay},   year={2023}, } ```</pre>