Found in Translation:
Learning Robust Joint Representations by Cyclic Translations Between Modalities

Paul Pu Liang
Joint work with: Hai Pham, Thomas Manzini, Louis-Philippe Morency, Barnabas Poczos
Robust Representation Learning

Speaker’s behaviors

“This is unbelievable!”

language

visual

acoustic

noisy/missing at test time

multimodal representation

prediction

Loud
Learning Joint Representations: 2 modalities

Traditional Methods

Language Modality

Today was a great day!

Joint Representation

Prediction

Visual Modality

Both modalities required at test time!
Sensitive to missing/noisy visual modality.
Learning Robust Joint Representations: 2 modalities

Bimodal Cyclic Translations

<table>
<thead>
<tr>
<th>Language Modality</th>
<th>Joint Representation</th>
<th>Visual Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Today was a great day!</em></td>
<td>forward</td>
<td>forward</td>
</tr>
<tr>
<td></td>
<td>backward</td>
<td>backward</td>
</tr>
</tbody>
</table>

Prediction

Only language modality required at test time! Robust to noisy/missing visual modalities.