The Amazing Mysteries of the Gutter: Drawing Inferences Between Panels in Comic Book Narratives

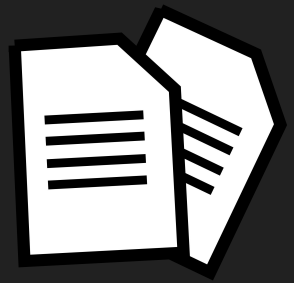
Mohit Iyyer Varun Manjunatha Anupam Guha, Yogarshi Vyas,
Jordan Boyd-Graber, Hal Daume III, Larry Davis

Presented by:
Leo Cao

Outlines

- Problem Definition
  - What is it & Why is it Hard
- Metrics (Loss Function)
- Data
- Methods
- Performance

x 5



x 2

The Problem is Hard

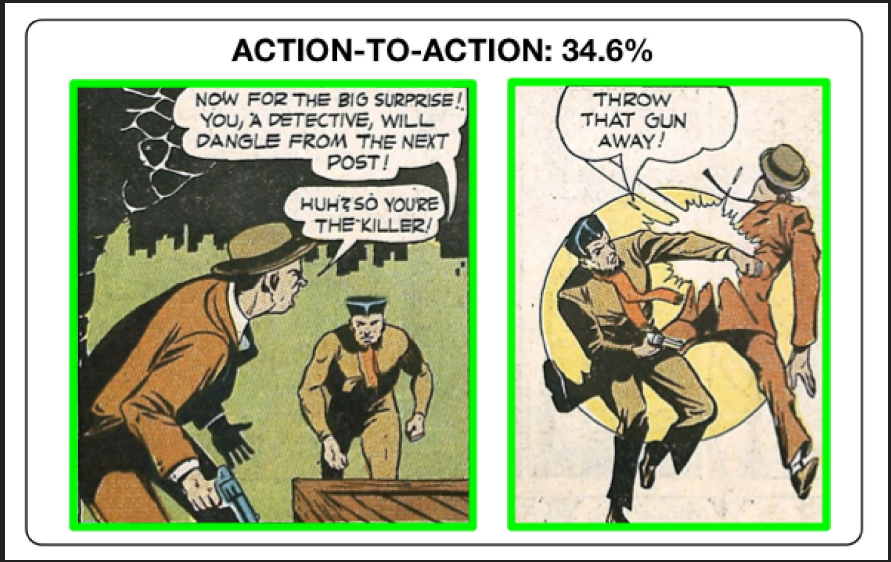
Comics = Stylized Artwork + Dialogue **(Vision + Language)**

- “Art” renders objects and concepts in many ways.

The Problem is Hard

Comics = Stylized Artwork + Dialogue **(Vision + Language)**

- Dialogues in comic often contain dialects and out-of-vocabulary tokens
- Judicious Omission



Metrics that Make it Easier (Avoiding Text Generation)

Loss Function: Cross-Entropy

Three Types of Tasks: (Back to this later)

- Text Cloze
- Visual Cloze
- Character Coherence

$$s = \mathrm{softmax}(A^T c),$$

Two Levels of Difficulty

- Easy: random candidates drawn from entire dataset
- Difficult: relevant candidates drawn from nearby sections

Data

Golden Age Comics:

- Small self-contained stories (no backstories, no Batman)
- Public domain data
- Plenty of them

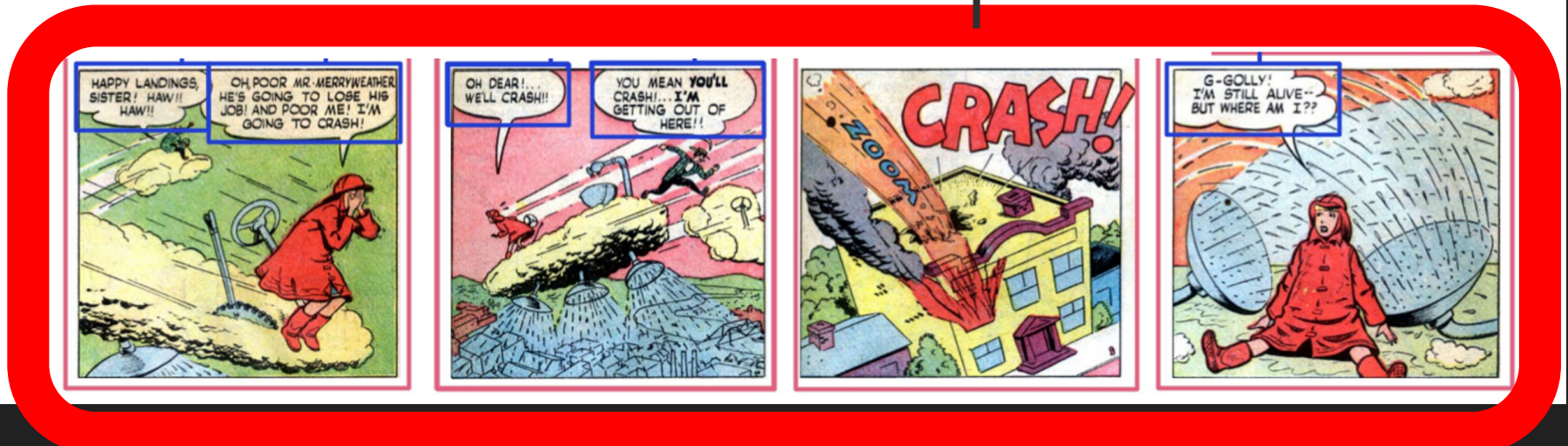
Processing:

- Panel Segmentation
- Text Segmentation
- OCR (Google Cloud Vision works the best, $3,000)

Text Colze

ALICE! I'VE BEEN LOOKING ALL OVER FOR YOU!

HIYA KID! ALL ALONE???



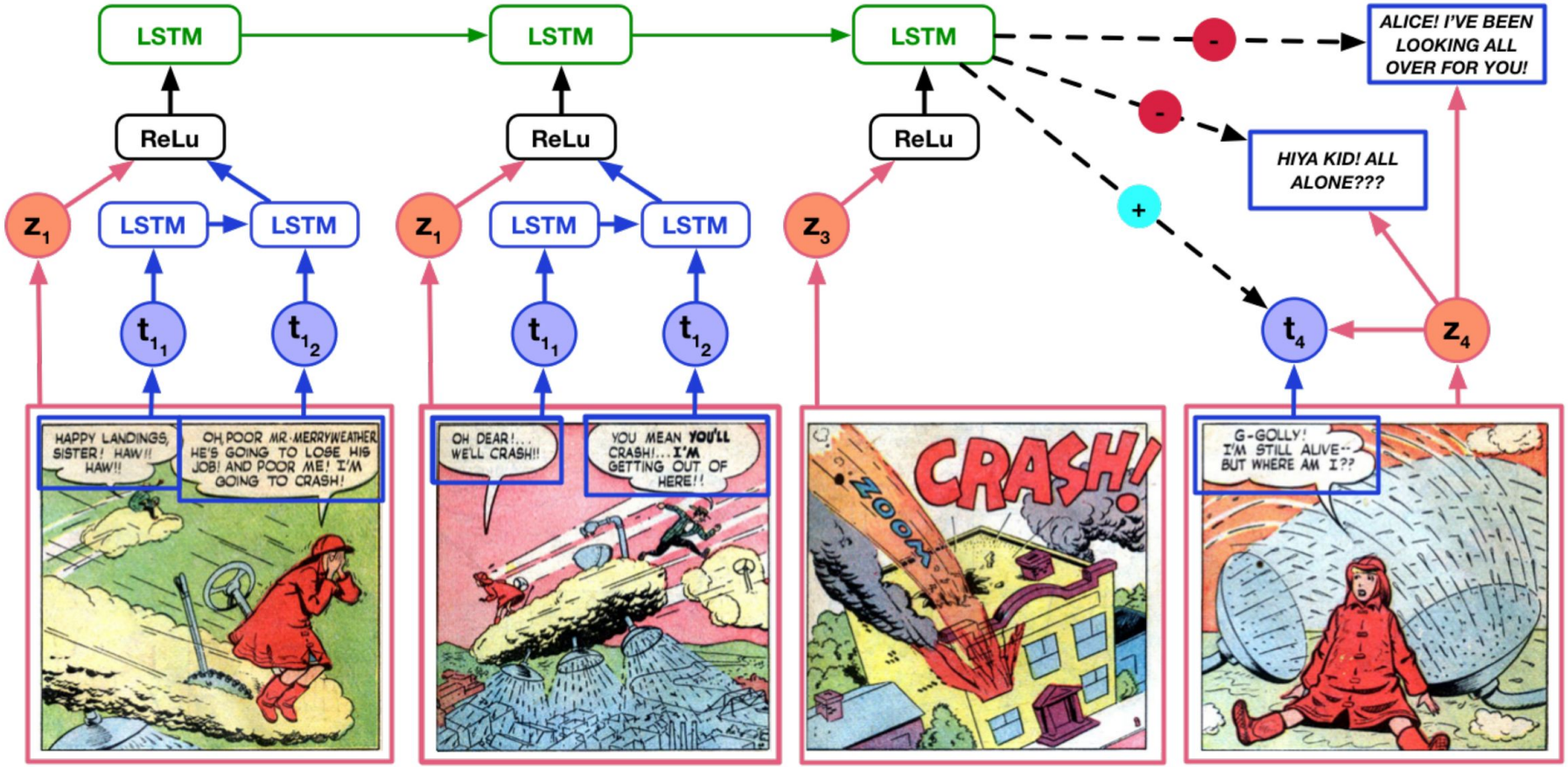
character coherence

1 2

2 1

visual cloze

Model


LSTM
LSTM
LSTM
ReLu
ReLu
ReLu
z_1
z_1
z_3
z_4
LSTM
LSTM
LSTM
LSTM
t_{1_1}
t_{1_2}
t_{1_1}
t_{1_2}
t_4
-
-
+
ALICE! I'VE BEEN LOOKING ALL OVER FOR YOU!
HIYA KID! ALL ALONE???
HAPPY LANDINGS, SISTER! HAW!! HAW!!
OH, POOR MR. MERRYWEATHER HE'S GOING TO LOSE HIS JOB! AND POOR ME! I'M GOING TO CRASH!
OH DEAR!... WE'LL CRASH!!
YOU MEAN YOU'LL CRASH!... I'M GETTING OUT OF HERE!!
ZOOM
CRASH!
G-GOLLY! I'M STILL ALIVE-- BUT WHERE AM I??

Results

Model	Text Cloze		Visual Cloze		Char. Coheren.
	easy	*hard*	*easy*	*hard*	
Random	33.3	33.3	33.3	33.3	50.0
Text-only	62.6	51.9	57.5	49.7	67.6
Image-only	59.8	48.7	82.3	**62.5**	67.1
NC-image-text	63.1	59.6	-	-	65.2
Image-text	**71.0**	**63.4**	**83.0**	61.4	**67.9**
Human	–	84	–	88	87

Questions?