

Aligning videos with recipes for video narration

Creating a dataset and Baseline

Video narration

Given a **silent video** and a **story**, the goal of video narration is to produce a narrated video where the video events are told in synchrony of the video, providing context to the video events to convey the focus message of the whole video.

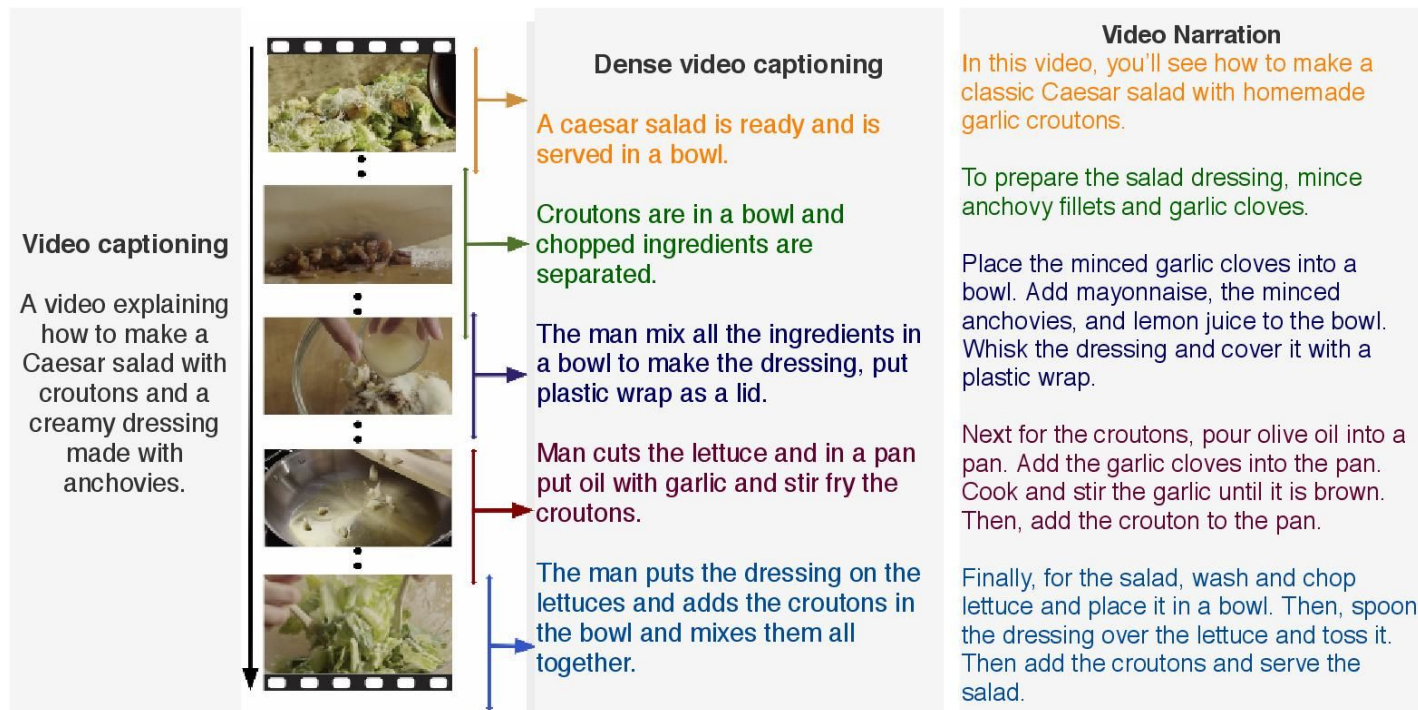


Fig: The frame examples and dense captioning annotations are of a sample from ActivityNet captions dataset [42].

Video narration is the transcript of the voice-over extracted from the original video.

Existing benchmarks

Existing benchmarks do not include needed elements to study video narration problem.

- The video.
- List of aligned instructions.
- Voice-overs transcripts.

Dataset	Duration	# Sample	Type of tasks	Recipe/ Instruction	Voice-over transcripts
MPII	9h48m	44	Cooking	<input type="checkbox"/>	X
YouCook	2h20m	88	Cooking	<input type="checkbox"/>	X
50Salads	5h20m	50	Cooking	<input type="checkbox"/>	X
Breakfast	77h	1989	Cooking	<input type="checkbox"/>	X
5 tasks	5h	150	Comprehensive	<input type="checkbox"/>	X
Ikea-FA	3h50m	101	Furniture assembling	<input type="checkbox"/>	X
YouCookII	176h	2000	Cooking	<input type="checkbox"/>	X
EPIC-Kitchen	55h	432	Cooking	<input type="checkbox"/>	X
COIN	476h38m	11827	Comprehensive	<input type="checkbox"/>	X

Humane-AI Micro-project Proposal

Main tasks:

1. **Clean** and extend an internally collected **dataset** and publish it in AI4EU platform.
 - **1058** videos, (a total duration of **69.1** hours)
 - average duration of 3.9 minutes per video.

Rich Banana Bread

Category: Bakery, Bread, Banana Bread **Meta-info**

Description: Sour cream guarantees a moist and tender loaf. And bananas are sliced instead of mashed in this recipe, giving a concentrated banana taste in every bite.


Rating: 6916 likes **Recipe**

Ingredients:

Name	Description
stick butter	0.5 cup butter, melted
white sugar	1 cup white sugar
eggs	2 eggs
vanilla extract	1 teaspoon vanilla extract
all-purpose flour	1.5 cups all-purpose flour
sour cream	0.5 cup sour cream
banana	2 medium bananas, sliced

Instructions: **1)** Preheat oven to 350 degrees F (175 degrees C). Grease a 9x5 inch loaf pan.
2) In a large bowl, stir together the melted butter and sugar. Add the eggs and vanilla, mix well. Combine the flour, baking soda and salt, stir into the butter mixture until smooth. Finally, fold in the sour cream, and bananas. Spread evenly into the prepared pan.
3) Bake at 350 degrees F (175 degrees C) for 60 minutes, or until a toothpick inserted into the center of the loaf comes out clean. Cool loaf in the pan for 10 minutes before removing to a wire rack to cool completely.

Voice-over transcript



[00:09 - 00:30] Allrecipes member Sherry were eggs that this recipe for Janet's Rich. Banana bread is one of the best she's ever found. She says the sour cream as a wonderfully moist quality to the bread and notes that she can't seem to make enough of it the first step to this delectable banana bread is to preheat the oven to 350 degrees then grease and 9 by 5 inch loaf pan.

[00:35 - 00:37] next peel the medium bananas.

[00:43 - 00:46] place the bananas in a bowl and mash them with a fork.

[00:55 - 00:58] In another small bowl combine one and a half cups of flour.

[01:05 - 01:12] one teaspoon of baking soda and half a teaspoon of salt.

[01:13 - 01:22] Use a fork to mix that together really well and set it aside now melt half a cup of butter. That's one stick and pour it into a large bowl.

[01:23 - 01:33] Add 1 cup of sugar and stir the mixture until it's well Blended.

[01:35 - 01:36] Add two eggs.

[01:42 - 01:46] And 1 teaspoon of vanilla extract and give it a good whisk to mix it well.

[01:53 - 01:56] Gently stir in the flour mixture until the batter is smooth.

[02:03 - 02:05] Then fold in half a cup of sour cream.

[02:23 - 02:25] And finally the mashed bananas.

[02:37 - 02:39] Spread the batter evenly into the prepared pan.

[02:45 - 02:48] bake the banana bread for that 60 minutes

[02:50 - 03:08] To check for doneness insert a toothpick into the center of the loaf. If it comes out clean the bread is done. If it's not quite clean continue the baking for another 5 to 10 minutes and check it again, cool the loaf in its pan for about 10 minutes then remove it to a cooling rack so it can finish cooling.

[03:08 - 03:20] Although it's really tempting to sample Janet's Rich banana bread hot from the oven the loaf will slice best if you let it cool completely. So try to exercise your willpower.

Video

Humane-AI Micro-project Proposal

Main tasks:

1. **Clean** and extend an internally collected **dataset** and publish it in AI4EU platform.
2. **Design**/reproduce a **baseline** method for video-text alignment which aligns instructions with the video.
 - **Video-text alignment:** Given a video and unordered list of instructions, localize where each instruction happened in the video.
 - Related works: [1, 2, 3]

[1] Anne Hendricks L, Wang O, Shechtman E, Sivic J, Darrell T, Russell B. **Localizing moments in video with natural language**. In Proceedings of the IEEE international conference on computer vision 2017 (pp. 5803-5812).

[2] Xu, Huijuan, Kun He, Bryan A. Plummer, Leonid Sigal, Stan Sclaroff, and Kate Saenko. "**Multilevel language and vision integration for text-to-clip retrieval**." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 9062-9069. 2019.

[3] Gavrilyuk K, Ghodrati A, Li Z, Snoek CG. **Actor and action video segmentation from a sentence**. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018 (pp. 5958-5966).

Collaboration and Contact

Collaboration:

- We are **looking for** collaboration in Humane-AI network to help us in advancing in this project.
- We are interested in collaborating with **NLP expertise**

Contact:

- James Crowley James.crowley@inria.fr
- Remi Ronfard Remi.ronfard@inria.fr
- Nachwa Baker Nachwa.aboubakr@inria.fr

Thank you for listening !