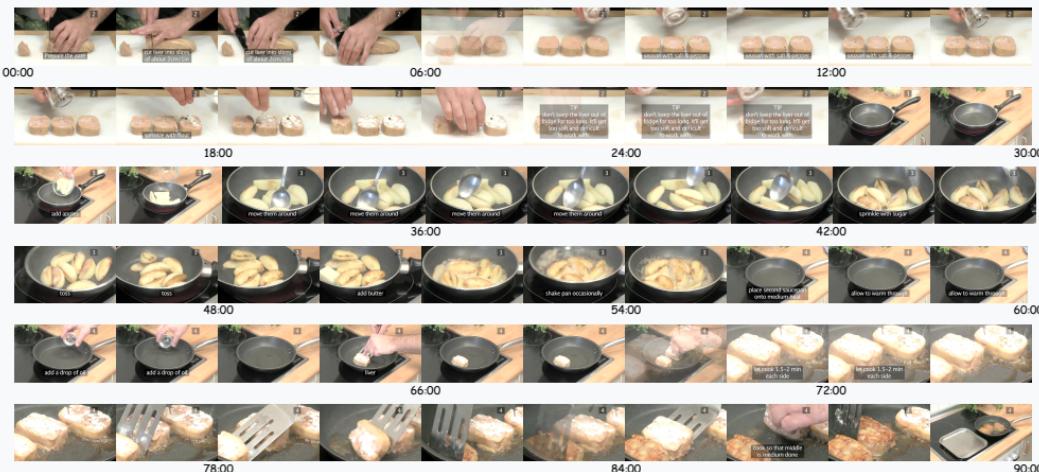


# Narration



Question: Describe various activity events occurring in the video in real time.

Ground Truth:

- 6.0s | cut the goose liver into small slices
- 12.0s | season them with salt and pepper
- 22.0s | sprinkle with flour on both sides
- 30.0s | heat the apple in a pan with some oil
- 48.0s | sprinkle some salt on top of the apples and add some butter
- 81.0s | cook the liver on another pan with oil

## ROMA (Ours)

- 7.0s | Cut the liver into slices.
- 13.0s | Season the liver with salt and pepper.
- 24.0s | Sprinkle flour on both sides of the liver.
- 42.0s | Add apples to the pan and move them around.
- 59.0s | Add butter to the pan and shake the pan occasionally.
- 71.0s | Place the liver in the pan and cook until browned.

## MMDuet

- 5.0s | A person is cutting a piece of meat into slices on a cutting board.
- 11.0s | sprinkle salt and pepper
- 16.0s | sprinkle flour on the meat
- 21.0s | sprinkle flour on the meat
- 27.0s | add seasoning and flour to the meat
- 34.0s | pour oil into the pan
- 40.0s | fry meat
- 46.0s | fry the meat
- 52.0s | fry the meat
- 58.0s | add butter
- 64.0s | fry the meat
- 70.0s | add the ingredients to the pan
- 83.0s | put the liver into the pan
- 87.0s | fry meat

## VideoLLM-Online

- 0.0s | (Video Time = 0s) Assistant: You cut the meat with a knife.