Using Inpainting to remove the Halo out of Formula 1 Onboard Visor Cam Footage

Sean O'Connor '26
CSCI 365: Image Processing and Analysis

May 7th, 2026

Bucknell
UNIVERSITY

F1 75

What is Formula 1?

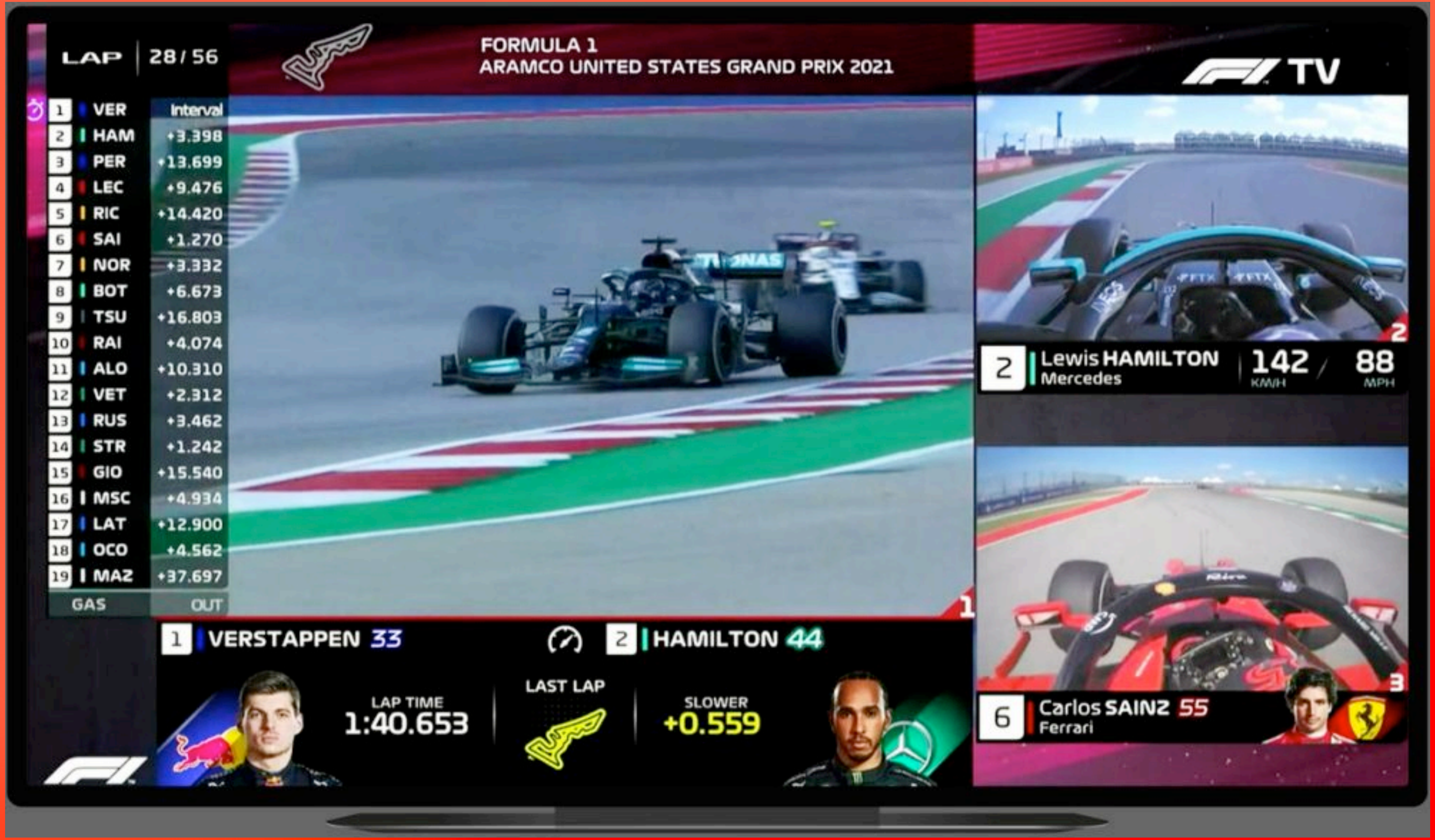
- Formula 1, known as the pinnacle of motorsport, is the highest caliber of open-wheel racing.
- 22 weekends a year, 20 drivers on 10 teams race at tracks around the world.

The Engineer's Sport

F1 is full of data- each car has 6 cameras, the tracks have ~20, each car produces over a terabyte of telemetry per race.

During a race, a viewer can select from over 25 live feeds, showing any number at one time.

Data is live-streamed to APIs such as FastF1, which will tell the public all of the car's data!



It's Really Fast!

Formula 1 cars can go up to 230 mph during a race.

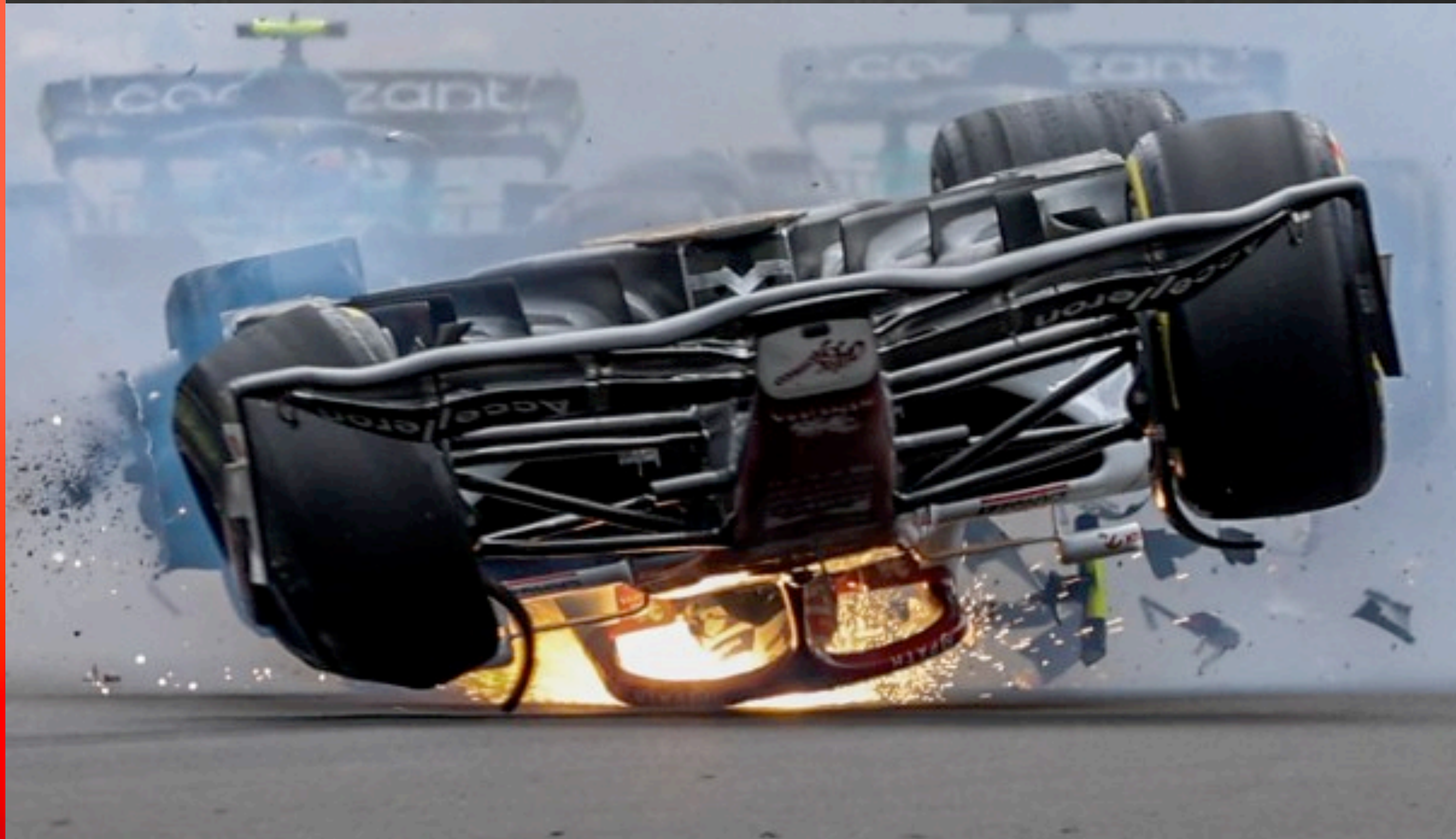
Crashes happen often at those speeds- drivers can take out multiple drivers if they just slightly miss the apex of a turn, or their wheels lock up.


SC

The Halo

Enter the Halo device. The halo is a driver crash-protection system, which consists of a curved bar placed above the driver's head to protect it from injury.

The halo saves lives!




MAX
VERSTAPPEN
RED BULL RACING
1
2


MAX
VERSTAPPEN
RED BULL RACING
1

How do we detect the halo?

The Halo moves!

Masking the Halo with Thresholding

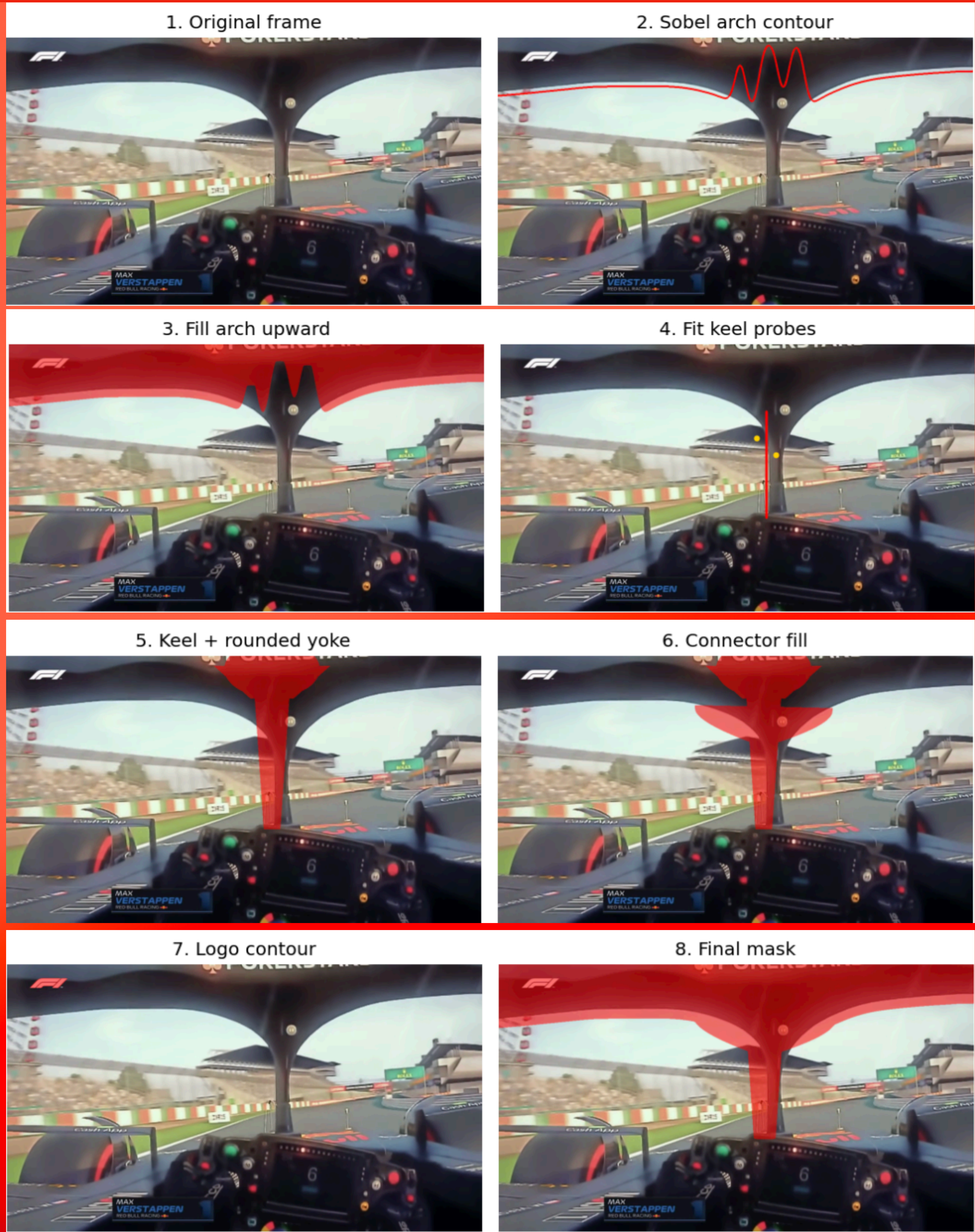
We can’t just use a static threshold mask- the halo angle moves with the visor.

Mask Detection: Geometry over Thresholding

So, we generate a mask based on the contours!

The Masking Pipeline

1. Detect the Halo arch edge with vertical Sobel gradients
2. Fill upward from the detected arch contour to the video border
3. Fit the moving keel from dark valley probe rows
4. Reject bad keel fits with outlier checks and temporal jump guards
5. Add rounded yoke + connector fill where the arch meets the keel
6. Add fixed F1 logo contour mask
7. Apply close / hole-fill / dilate cleanup



Now, let’s remove the Halo.

We’ll use inpainting.

What is it: Reconstructing missing or hidden regions of an image so the result looks like it was never damaged.

Three methods of note:

- **PatchMatch**: Find the best matching patch from elsewhere in the same image and copy it in.
- **LaMa (Neural, 2022):** A CNN trained on 100k natural scenes. Uses Fast Fourier Convolutions to group and extend areas of an image.
- **Temporal Propagation:** Instead of filling from the current frame alone, borrow the previous clean frame and warp it forward in time using optical flow.

PatchMatch


MAX
VERSTAPPEN
RED BULL RACING
MAX
VERSTAPPEN
RED BULL RACING

Replacing the Halo with Large Mask Inpainting

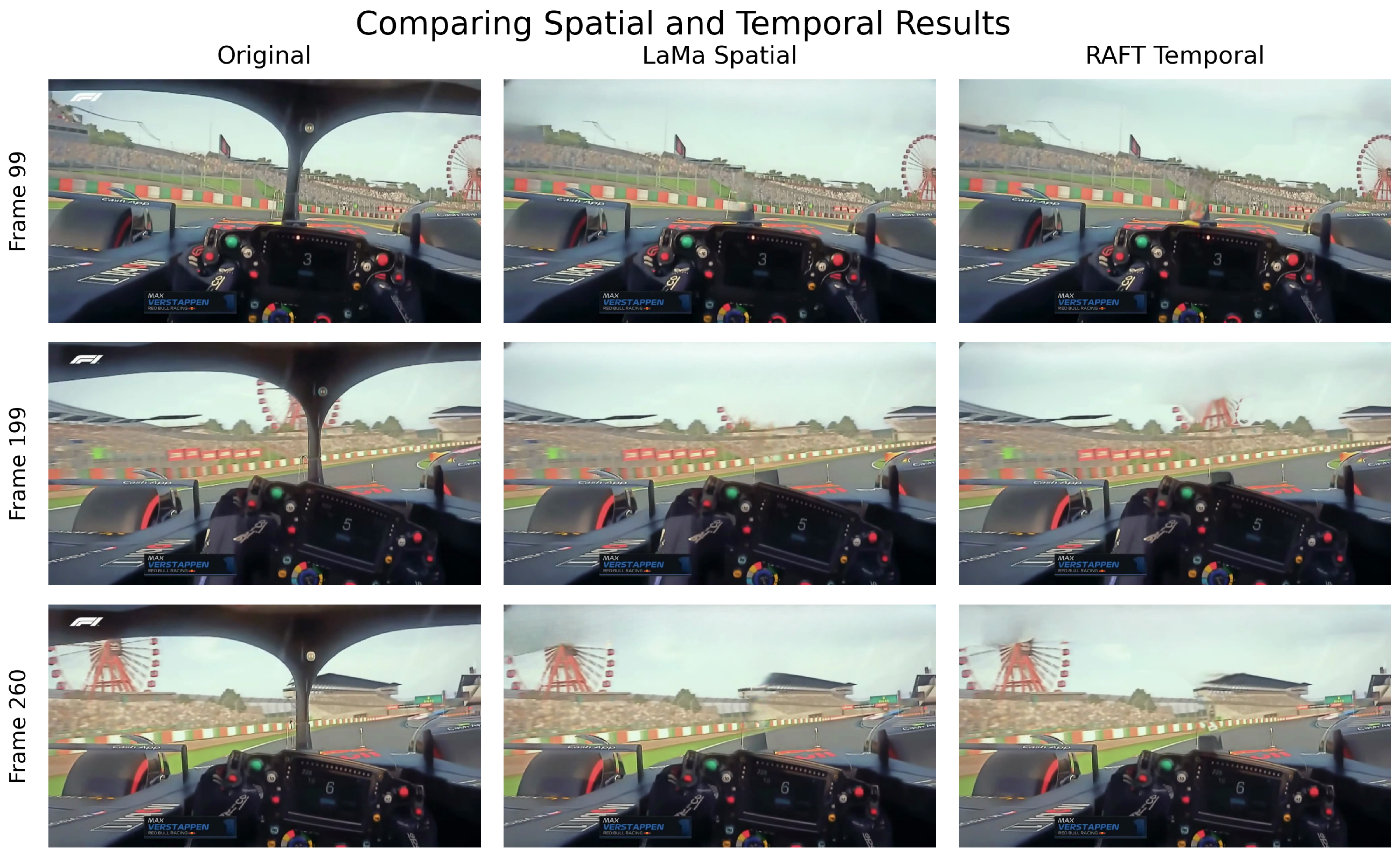
- Pretrained neural inpainting model for large missing regions
- Takes the original frame and binary Halo mask as input
- Uses global image context to synthesize plausible replacement pixels
- Runs independently on each frame
- Produces the cleanest single-frame results in the pipeline

RAFT Temporal Propagation

- Uses optical flow to estimate motion between frames
- Warps the previous clean frame into the current frame
- Blends the warped result inside the Halo mask
- Goal: reduce frame-to-frame flicker from independent LaMa outputs
- Works best when the previous inpainted frame is already clean

Comparing Methods

- LaMa produces better individual frames, RAFT improves temporal consistency.
- RAFT can also carry artifacts forward
- Best method depends on spatial quality vs. temporal stability



Comparing Methods: Video

Original

Spatial (LaMa)

Temporal (RAFT)

Limitations & Next Steps

- **Keel occlusion:** When hands or the steering wheel fully cover the keel, the detector falls back to the previous frame's center. Stable, but the inpainter is now filling a large contiguous region without any current-frame evidence of where the keel actually is. This causes occasional smearing at the keel base.
- **Arch/keel junction:** The convergence point where the keel meets the arch is the most failure-prone pixel region. A small mask gap here leaves a visible dark edge in the inpainted output. The yoke fill and connector line improve it, but it remains sensitive.
- **RAFT error accumulation:** The temporal chain resets to the LaMa output every 60 frames. Without resets, warp drift compounds: a slightly wrong fill at frame 60 becomes increasingly wrong by frame 120. This tradeoff is baked into the design, not a bug.

Takeaways

- **Large-mask neural inpainting outperforms PatchMatch by a wide margin:** not because PatchMatch is a bad algorithm, but because this is the wrong problem for it. Any technique that copies pixels from the same frame fails the moment the entire content class (sky) is inside the mask.
- **Per-frame detection beats tracking for occluded structures:** Lucas-Kanade is great when the thing you're tracking is always visible. The keel is not. Independent detection per frame means you're always working from real image evidence rather than an accumulated guess.
- **Temporal consistency doesn't come for free:** LaMa alone produces independent, high-quality fills, but they flicker. Adding RAFT doesn't magically fix that; it trades some spatial quality for reduced shimmer, and it can propagate artifacts.

Thank you!

Any questions?