



AiA Art News-service



## Mona Lisa frown: Machine learning brings old paintings and photos to life

[Devin Coldewey@techcrunch](#) / 17 hours ago

Comment



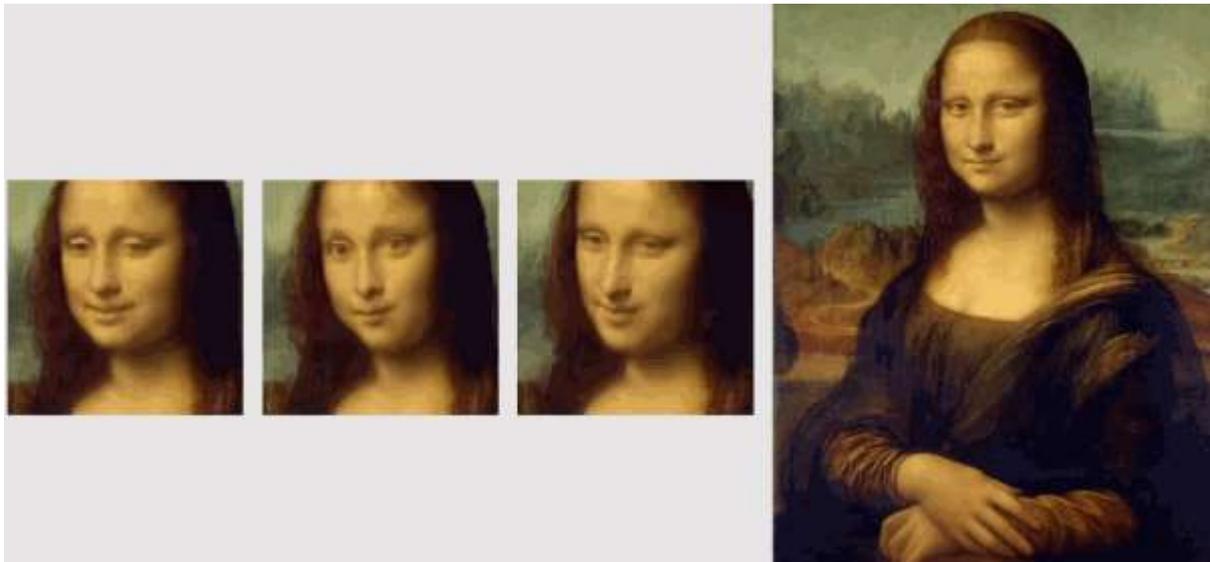
Machine learning researchers have produced a system that can recreate lifelike motion from just a single frame of a person's face, opening up the possibility of animating not just photos but also paintings. It's not perfect, but when it works, it is — like much AI work these days — eerie and fascinating.

The model is documented in a paper published by Samsung AI Center, which [you can read here on Arxiv](#). It's a new method of applying facial landmarks on a source face — any talking head will do — to the facial data of a target face, making the target face do what the source face does.

This in itself isn't new — it's part of the whole synthetic imagery issue confronting the AI world right now ([we had an interesting discussion about this recently at our Robotics + AI event in Berkeley](#)). We can already make a face in one video reflect the face in another in terms of what the person is saying or where they're looking. But most of these models require a considerable amount of data, for instance a minute or two of video to analyze.

The new paper by Samsung's Moscow-based researchers, however, shows that using only a single image of a person's face, a video can be generated of that face turning, speaking and making ordinary expressions — with convincing, though far from flawless, fidelity.

It does this by frontloading the facial landmark identification process with a huge amount of data, making the model highly efficient at finding the parts of the target face that correspond to the source. The more data it has, the better, but it can do it with one image — called single-shot learning — and get away with it. That's what makes it possible to take a picture of Einstein or Marilyn Monroe, or even the Mona Lisa, and make it move and speak like a real person.



*In this example, the Mona Lisa is animated using three different source videos, which as you can see produce very different results, both in facial structure and behavior.*

It's also using what's called a Generative Adversarial Network, which essentially pits two models against one another, one trying to fool the other into thinking what it creates is "real." By these means the results meet a certain level of realism set by the creators — the "discriminator" model has to be, say, 90% sure this is a human face for the process to continue.

In the other examples provided by the researchers, the quality and obviousness of the fake talking head varies widely. Some, which attempt to replicate a person whose image was taken from cable news, also recreate the news ticker shown at the bottom of the image, filling it with gibberish. And the usual smears and weird artifacts are omnipresent if you know what to look for.

That said, it's remarkable that it works as well as it does. Note, however, that this only works on the face and upper torso — you couldn't make the Mona Lisa snap her fingers or dance. Not yet, anyway.

