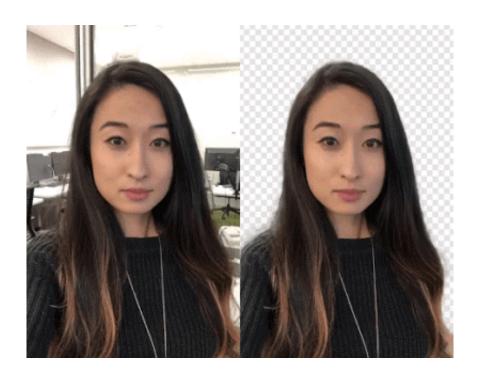# Background-swap technology suitable for mobile phones announced

March 4 2018, by Nancy Owano

Credit: Google

Video creators: Want to swap backgrounds? Knock yourselves out. Google researchers have been working on a way to let you swap out your video backgrounds using a neural network—no green screen required. It's rolling out to YouTube Stories on mobile in a limited fashion, said *TechCrunch*.

John Anon, *Android Headlines*, said YouTube creators can change the

background to create more engaging videos.

Valentin Bazarevsky and Andrei Tkachenka, software engineers, Google Research, made the announcement, titled "Mobile Real-time Video Segmentation."

Video content creators know that a scene's background can be separated from the background treated as two different layers. The maneuver is done to achieve a mood, or insert a fun location or punch up the impact of the message.

The operation, said the two on the Google Research site, is "a time-consuming manual process (e.g. an artist rotoscoping every frame) or requires a studio environment with a green screen for real-time background removal (a technique referred to as chroma keying)."

Translation: Hillary Grigonis in *Digital Trends* said, "Replacing the background on a video typically requires advanced desktop software and plenty of free time, or a full-fledged studio with a green screen."

Now the two have announced a new technique, and it will work on mobile phones.

Their technique will enable creators to replace and modify backgrounds without specialized equipment.

They called it YouTube's new lightweight video format, designed specifically for YouTube creators.

They issued a March 1 announcement of a "precise, real-time, on-device mobile video segmentation to the YouTube app by integrating this technology into stories."

How did they do this? Anon said "the crux of it all is machine [learning](#)."

Bazarevsky and Tkachenka said they leveraged "machine learning to solve a semantic segmentation task using convolutional neural networks."

Translation: "Google is developing an artificial intelligence alternative that [works](#) in real time, from a smartphone camera," Grigonis wrote.

The two engineers described an architecture and training procedure suitable for mobile phones. They kept in mind that "A mobile solution should be lightweight and run at least 10-30 times faster than existing state-of-the-art photo segmentation models."

As for a dataset, they "annotated tens of thousands of images." These captured a wide spectrum of foreground [poses](#) and background settings.

"With that data set, the group trained the program to separate the background from the foreground," said Grigonis.

Devin Coldewey in *TechCrunch*: "The network learned to pick out the common features of a [head](#) and shoulders, and a series of optimizations lowered the amount of data it needed to crunch in order to do so."

Digital Trends explained how it works: "Once the software masks out the background on the first image, the program uses that same mask to predict the background in the next frame. When that next frame has only minor adjustments from the first...the program will make small adjustments to the mask. When the next frame is much different from the last ...the software will discard that mask prediction entirely and create a new mask."

One end result of their work, as said on the Google Research blog, is that "our network runs remarkably fast on mobile devices, achieving 100+

FPS on iPhone 7 and 40+ FPS on Pixel 2 with high accuracy (realizing 94.8% IOU on our validation dataset), delivering a variety of smooth running and responsive effects in YouTube stories."

What's next?

It is in limited beta. "Our immediate goal is to use the limited rollout in YouTube stories to test our technology on this first set of effects. As we improve and expand our segmentation technology to more labels, we plan to integrate it into Google's broader Augmented Reality services."