

How AI is changing photography

If you're wondering how good your next phone's camera is going to be, it'd be wise to pay attention to what the manufacturer has to say about AI. Beyond the hype and bluster, the technology has enabled staggering advances in photography over the past couple of years, and there's no reason to think that progress will slow down.

There are still a lot of gimmicks around, to be sure. But the most impressive recent advancements in photography have taken place at the software and silicon level rather than the sensor or lens – and that's largely thanks to AI giving cameras a better understanding of what they're looking at.

Google Photos provided a clear demonstration of how powerful a mix AI and photography would be when the app launched in 2015. Prior to then, the search giant had been using machine learning to categorize images in Google+ for years, but the launch of its Photos app included consumer-facing AI features that would have been unimaginable to most. Users' disorganized libraries of thousands of untagged photos were transformed into searchable databases overnight.

Suddenly, or so it seemed, Google knew what your cat looked like.

Photo by James Bareham / The Verge

Google built on the previous work of a 2013 acquisition, DNNresearch, by setting up a deep neural network trained on data that had been labeled by humans. This is called supervised learning; the process involves training the network on millions of images so that it can look for visual clues at the pixel level to help identify the category. Over time, the

algorithm gets better and better at recognizing, say, a panda, because it contains the patterns used to correctly identify pandas in the past. It learns where the black fur and white fur tend to be in relation to one another, and how it differs from the hide of a Holstein cow, for example. With further training, it becomes possible to search for more abstract terms such as “animal” or “breakfast,” which may not have common visual indicators but are still immediately obvious to humans.

It takes a lot of time and processing power to train an algorithm like this, but after the data centers have done their thing, it can be run on low-powered mobile devices without much trouble. The heavy-lifting work has already been done, so once your photos are uploaded to the cloud, Google can use its model to analyze and label the whole library. About a year after Google Photos was launched, Apple announced a photo search feature that was similarly trained on a neural network, but as part of the company’s commitment to privacy the actual categorization is performed on each device’s processor separately without sending the data. This usually takes a day or two and happens in the background following setup.

Intelligent photo management software is one thing, but AI and machine learning are arguably having a bigger impact on how images are captured in the first place. Yes, lenses continue to get a little faster and sensors can always get a little bigger, but we’re already pushing at the limitations of physics when it comes to cramming optical systems into slim mobile devices. Nevertheless, it’s not uncommon these days for phones to take better photos in some situations than a lot of dedicated camera gear, at least before post-processing. That’s because traditional cameras can’t compete on another category of hardware that’s just as profound for photography: the systems-on-chip that contain a CPU, an image signal processor, and, increasingly, a neural processing unit (NPU).

This is the hardware leveraged in what's come to be known as computational photography, a broad term that covers everything from the fake depth-of-field effects in phones' portrait modes to the algorithms that help drive the Google Pixel's incredible image quality. Not all computational photography involves AI, but AI is certainly a major component of it.

Apple makes use of this tech to drive its dual-camera phones' portrait mode. The iPhone's image signal processor uses machine learning techniques to recognize people with one camera, while the second camera creates a depth map to help isolate the subject and blur the background. The ability to recognize people through machine learning wasn't new when this feature debuted in 2016, as it's what photo organization software was already doing. But to manage it in real time at the speed required for a smartphone camera was a breakthrough.

Google remains the obvious leader in this field, however, with the superb results produced by all three generations of Pixel as the most compelling evidence. HDR+, the default shooting mode, uses a complex algorithm that merges several underexposed frames into one and, as Google's computational photography lead Marc Levoy has noted to *The Verge*, machine learning means the system only gets better with time. Google has trained its AI on a huge dataset of labeled photos, as with the Google Photos software, and this further aids the camera with exposure. The Pixel 2, in particular, produced such an impressive level of baseline image quality that some of us at *The Verge* have been more than comfortable using it for professional work on this site.

Google's Night Sight is a stunning advertisement for the role of software in photography

But Google's advantage has never seemed so stark as it did a couple of months ago with the launch of Night Sight. The new

Pixel feature stitches long exposures together and uses a machine learning algorithm to calculate more accurate white balance and colors, with frankly astonishing results. The feature works best on the Pixel 3, because the algorithms were designed with the most recent hardware in mind, but Google made it available for all Pixel phones – even the original, which lacks optical image stabilization – and it's a stunning advertisement for how software is now more important than camera hardware when it comes to mobile photography.

That said, there is still room for hardware to make a difference, particularly when it's backed by AI. Honor's new View 20 phone, along with parent company Huawei's Nova 4, are the first to use the Sony IMX586 image sensor. It's a larger sensor than most competitors and, at 48 megapixels, represents the highest resolution yet seen on any phone. But that still means cramming a lot of tiny pixels into a tiny space, which tends to be problematic for image quality. In my View 20 tests, however, Honor's "AI Ultra Clarity" mode excels at making the most of the resolution, descrambling the sensor's unusual color filter to unlock extra detail. This results in huge photographs that you can zoom into for days.

Image signal processors have been important to phone camera performance for a while, but it looks likely that NPUs will take on a larger role as computational photography advances. Huawei was the first company to announce a system-on-chip with dedicated AI hardware, the Kirin 970, although Apple's A11 Bionic ended up reaching consumers first. Qualcomm, the biggest supplier of Android processors worldwide, hasn't made machine learning a major focus yet, but Google has developed its own chip called the Pixel Visual Core to help with AI-related imaging tasks. The latest Apple A12 Bionic, meanwhile, has an eight-core neural engine that can run tasks in Core ML, Apple's machine learning framework, up to nine times faster

than the A11, and for the first time it's directly linked to the image processor. Apple says this gives the camera a better understanding of the focal plane, for example, helping generate more realistic depth of field.

The camera is an essential feature of any phone, and AI is our best shot at improving it

This kind of hardware will be increasingly important for efficient and performant on-device machine learning, which has an exceptionally high ceiling in terms of its demands on the processor. Remember, the kind of algorithms that power Google Photos were trained on huge, powerful computers with beefy GPUs and tensor cores before being set loose on your photo library. Much of that work can be done "in advance," so to speak, but the ability to carry out machine learning calculations on a mobile device in real time remains cutting edge.

Google has shown some impressive work that could reduce the processing burden, while neural engines are getting faster by the year. But even at this early stage of computational photography, there are real benefits to be found from phone cameras that have been designed around machine learning. In fact, out of all the possibilities and applications raised by the AI hype wave of the past few years, the area with the most practical use today is arguably photography. The camera is an essential feature of any phone, and AI is our best shot at improving it.

<https://www.theverge.com/2019/1/31/18203363/ai-artificial-intelligence-photography-google-photos-apple-huawei>