THE COMPUTER ATE MY PHOTOS

Artificial Intelligence and the Future of Photography

Aaron Hockley

The Computer Ate My Photos

Artificial Intelligence and the Future of Photography

Aaron Hockley

The Computer Ate My Photos: Artificial Intelligence and the Future of Photography

Independently published by Aaron Hockley
13504 NE 84th St, Suite 103-101, Vancouver, WA 98682
techphotoguy.com

Aaron Hockley, also known as the Tech Photo Guy, brings together his background in technology and photography to write and speak about the intersection of those worlds. He lives in Vancouver, Washington with his family and can be found on the web at techphotoguy.com

For updates on the material in this book, information about the latest developments in photography technology, and insight into the evolving modern photo industry, get more information about what's coming next: next.techphotoguy.com

This book is dedicated to everyone who has moved past arguing about whether the artistic or technical aspects of photography are more important and instead is embracing the union of the two.

1. Welcome to the Third Shift

I got married in the year 2000.

It was a fairly traditional wedding. As such, we engaged in the fairly traditional activity of hiring a professional wedding photographer. We met with several folks in the area and made our selection. Digital photography was new at the time and while some folks were experimenting, all the "real" photographers were still working with film. The expectation was that serious photographers used film and delivered prints, and you hired a professional because only professionals had the gear, skill, and resources to make high-quality images.

My, how things have changed.

The world of popular photography has seen significant changes that have altered the photographic landscape for casual photo hobbyists, advanced amateurs, and professional photographers alike. As we look toward the future, with computers and photography becoming ever-intertwined, we are in the early stages of a third big shift in the photographic worlds related to these two subjects.

Let's explore the previous two shifts and how they've set us up for today.

Hello Internet

For all but the youngest of this book's readers, we remember the photography of our childhood consisting of a variety of print products. Our dad had pictures of us and our siblings in his wallet. Images from the family road trip were printed and stored in plastic pages of a photo album that lived on a bookshelf or the coffee table. At family events, a group picture was taken and copies were shared (often sent through the US mail) to relatives everywhere, whether they'd been in attendance or were in a "wish you were here" situation. Our latest school portrait would hang in the hallway or be placed in a frame on a shelf.

Casual distribution of photographs involved making physical prints and then getting those tangible items to the recipient. It worked well for folks in the same house or town but was challenging for those across the country or around the world. If my parents photographed my sister making an especially big mess of her dinner as a baby, it was a bit of work to share those images with distant relatives.

And then the internet happened.

In the 1990s, proprietary online services such as Prodigy, CompuServe, and America Online gave way to the open internet, and folks in the western world got connected. This access originally came in the form of (relatively slow) dial-up modem access. Who among us doesn't wax nostalgic for the screeching modem tones? Have we forgotten the experience of being disconnected because someone in another part of the house picked up the telephone? The phone-based dial-up modems gave way to DSL and then cable modems as broadband internet access became part of mainstream life.

The Pew Research Center first started measuring United States internet use in 2000, when it found that half of all adults regularly used the internet[1]. That number grew to 90% in 2019 (with 73% on broadband), and I'd wager that for folks who are seriously into photography, the number is *really* close to 100%.

The first big shift in modern photography came when mainstream internet usage made it easy to distribute our pictures.

Even before digital cameras arrived on the scene, photography enthusiasts used scanners to digitize their images onto the computer, and it was then possible to send them anywhere in the world via email. Message boards and early websites made it easy to share work with a broader (and more public) audience.

In the early 2000s, entire communities and companies sprung up around the new world of online photo sharing. Beyond sharing our work with family and friends, these online systems allowed us to connect with other photographers. Flickr was the largest of these communities; there was a point where nearly *anyone* into photography in any way had a Flickr account and shared their work on the service.

Ten years prior, casual image sharing meant making prints, putting them in an envelope, and waiting for the postal

service. Thanks to the internet, it could now be done in a couple of minutes with a few clicks of the mouse.

Digital Cameras Everywhere

If we go back thirty years, one of the things that helped set professional photographers (and the serious hobbyists) apart from average folks was that making great photos required some fairly exclusive gear. Options were limited, and if one wanted to make the best images, it required a hefty investment.

In short, truly serious camera gear wasn't easily accessible to a person who wasn't into photography as their full-time profession.

The first digital cameras that made decent photos were in the same price category, making them only accessible to professionals or those with a *lot* of disposable income for a hobby. They weren't accessible to everyday folks. But like everything in the world of computing, the cost of the technology dropped over time, and eventually we hit a breakthrough in 2003 when Canon released its model 300D, known in the United States as the Digital Rebel. The release of the original Digital Rebel[2] marked the turning point in making high-quality, interchangeable-lens digital cameras accessible to serious photographic hobbyists. For less than $1000, we had a decent digital SLR available that could use Canon's existing lineup of lenses and accessories. After this camera from Canon, we quickly saw similar models show up from Nikon; these two companies' DSLRs gained a lot of popularity within a few years.

In addition to being affordable, the other big impact with the prosumer DSLRs from Canon and Nikon was that they made images of sufficient quality for most uses. Sure, there were still cases where one needed heftier gear, but for casual family portraits and vacation photos and children's sporting events, these cameras were quite capable.

The second big shift in photography was digital cameras becoming good enough (and affordable enough) that they reached widespread adoption among photo enthusiasts.

Folks who wanted to make great pictures no longer had equipment as a barrier. Professionals who used to use their (expensive and hard to learn) equipment as a distinguishing factor for their businesses started finding that their clients sometimes had gear that was as good or better than theirs.

Technical Plus Creative Equals Photography

Success in photography has always depended on a combination of factors: technical ability, creative vision, and the skill to turn that vision into an impactful photograph through the use of cameras, lights, and other equipment.

Whether we're talking about the "olden days" of film and flash powder or modern digital cameras with remotely-triggered strobes or LED lights, a photographer has needed some level of technical understanding. They needed to understand the technical operation of their gear. They needed to understand how the various settings on their camera and lights worked together to make an exposure.

- How does the shutter speed relate to the aperture?
- How does ISO factor into things?
- What's a lighting ratio and how do I adjust it?

None of these were insurmountable obstacles, but each required bit of learning and yes, when desperate, reading the manual. Even if one had a brilliant creative vision, some technical know-how was needed to execute that vision.

That vision is the other major factor needed by a photographer. Without some understanding of aesthetics, color theory, composition, and other creative elements, the most we can hope for is a technically correct (but boring) photograph. Even the most technically adept photographer may need to invest some time into understanding the creative aspects of visual art (many of which apply more broadly than just the photographic world).

After the shift in casual distribution from print to the internet, and the shift from complicated and expensive film cameras to affordable and accessible digital ones, where does that leave us with respect to the technical and creative factors?

Does the ever-more-computerized photography world affect the technical and creative factors? Most definitely. Technical advances are making new things possible and are flattening the learning curve for existing possibilities. While a computer can't fully replace the creative brain, it can certainly help execute a creative vision and might just be able to make the creative result easier to attain.

We'll take a more in-depth look at how computational advances are helping both the technical and creative worlds in the coming pages.

The Third Shift

The third major shift in modern photography is already underway.

This third shift is fueled by the rise of computational photography and artificial intelligence to reshape how we capture and process our images.

The third major shift in modern photography is a shift where photographers spend less time telling their cameras and computers how to make an image and instead allow the "smart" devices to make an increasing number of technical and creative decisions.

As we look at AI and photography overall, these advances aren't just about capturing still photos, editing, and sharing them. We also need to look at the intersections with video, voice, creativity, metadata, and the business aspects of the industry.

Folks who sell for a living (including photographers) are likely familiar with discussions around features versus benefits. Companies often talk about features, whereas customers generally care more about benefits. Consider:

- This lens is f/1.2 (feature) vs. this lens can create beautiful blurry backgrounds to isolate a subject (benefit)
- This camera features the new Digiwhiz 4 image processor (feature) vs. this camera's new image processor allows photos to be saved faster to your memory card (benefit)
- This software features integrated cloud services (feature) vs. this software allows you to easily synchronize your image edits between computers and mobile devices (benefit)

Customers want the benefits; features are the way we get there. As we spend time looking at artificial intelligence and how it affects cameras, lighting, software, and other parts of the photographic world, we will look at both features and benefits. Whether it's used embedded into hardware or part of a software package, AI is a feature and it can bring many benefits.

We're going to look at the intersection of artificial intelligence and photography in four groups of information:

- The basics of AI, machine learning, computational photography, and what we should understand about the technology and the terminology in general.
- How AI is (and will be) used by our cameras, smartphones, and software.
- How AI developments are shifting what we know and think about metadata, privacy, and our very definitions of photography itself.
- What an AI-powered photographic future looks like, and what it means for photographers, media companies, and society.

In the coming pages, you'll learn how artificial intelligence is prevalent throughout the photo world and we'll develop an eye to think about the future. There have likely been interesting developments in the AI photo world since this book was published, even if you're an early reader. Don't worry; I have you covered. You can go to next.techphotoguy.com to find out what's new and stay current.

Let's dive in…

1. https://www.pewinternet.org/fact-sheet/internet-broadband/

2. There have been several subsequent “Digital Rebel” model prosumer cameras.

2. Technical Plus Creative Plus Computers

Let’s explore the technical and creative mixture mentioned previously, and look at how computational advances (including artificial intelligence and the cloud) are going to further both areas of the photo world.

Computers and the Technical Realm

For the technical factor, as we look at modern photography and computational advances, we see two big impacts.

First, new things become possible that previously couldn’t be done. This might mean increased resolution of our images due to advances in image sensors. Megapixel count used to be the main statistic touted for new digital cameras, but eventually, we realized that the number of pixels isn’t the sole factor in image quality. That being said, higher resolution does allow for greater flexibility in how we use our images, or for the ability to make significant crops in post-production while still retaining useable image quality. Other technical advances might mean better low-light performance of the camera on our smartphone or DSLR due to new software algorithms that weren’t possible a few years ago. Unsurprisingly, one of the big factors in image quality is the size of the camera’s sensor. This is why full-frame[3] cameras typically deliver a higher-quality image than smaller cameras, and why early cell phone cameras were pretty crappy… their sensors were tiny! The sensors inside your current iPhone or Samsung smartphone aren’t much bigger, but they’ve gotten much better and are paired with some incredible image capture software included as part of the iOS and Android operating systems. Each year as new smartphones are released, new things become possible.

The other technical impact is that things that *were* previously possible can be made easier via technical advances. Perhaps something that previously required configuring many settings could now be accomplished by pushing a single button, with software figuring out the optimal situation.

One big area worthy of discussion is the ever-so-nebulous “cloud” that we hear about when it comes to modern computing.

The Cloud as Many Computers

There’s nothing magical about the cloud… ultimately it’s just a bunch of computers and storage run by various companies that run such computers and storage as part of their business. In 2021 there are a small handful of companies that are major players in cloud computing infrastructure. While other companies offer cloud services, ultimately it usually boils down to servers and storage managed by Amazon, Microsoft, Google, or Apple.

If we look at how cloud computing intersects with photography, we're generally talking about two uses.

The first is using the power of many computing devices to work through the artificial intelligence and machine learning tasks associated with intelligent photography software. How does Google learn to recognize objects in Google Photos? Though its cloud computing farms. When we speak of technical possibilities that weren't possible several years ago, these are sometimes because we can now aggregate the computing power across the many servers in the cloud. If one computer can crunch numbers at a given rate, many of those computers can crunch numbers at a much faster rate.

The second area where photographers use the cloud is when it comes to cloud storage, which usually means online backup storage through various vendors. This could mean a photography-specific service such as SmugMug, which includes unlimited cloud storage as part of all of its service plans. Or it could mean a more generalized backup solution such as those provided by BackBlaze, which can automatically back up a computer and all of its data to the company's online storage. For folks who want to have a bit more direct involvement in their cloud storage solution, consumers can purchase space directly from a service such as Amazon S3 or Amazon Glacier, then use a client application on their computer to back up data directly to the cloud service.

With cloud storage and backup, we're not relying on the fact that there can be a lot of different computing processors at work, but rather the fact that a company can have massive amounts of reliable online storage and make it available for backup and other purposes.

A Word on Privacy…

If we look at the ideas of uploading our images to servers and disk drives controlled by random internet companies, it's reasonable to also look at the privacy concerns that might surround that practice. Our photos often contain personal or sensitive moments, and we may not want those images to be seen by employees at Amazon, or Google, or Apple, let alone the entire world if there were some sort of issue and they became public.

Ultimately if we're looking for assurances around privacy, nearly all companies will say words that state they value your privacy and will take care to protect it. Unless one is to have a blanket trust (or distrust) of all companies regardless of other factors, we need something more substantial upon which to base our decisions about which companies do we trust, and which might be less trustworthy and lead us to exercise more caution when uploading our photography. Consider three main criteria to evaluate a company's trustworthiness when it comes to privacy:

- **Does the company have a track record?** We might not be able to judge a brand-new startup, but when we're talking about the biggest players in the cloud photo space, we're often talking about industry veterans. Assuming the company has been around for a while, we look at the next question.
- **What is the company's history when it comes to privacy?** Actions speak far louder than words. If a company has a history of repeated privacy breaches or situations where they misled the public about their use of individuals' data, a wise individual would be skeptical to trust that company's claims of privacy in the future. We have no stronger example of this than Facebook, which has had literally dozens of instances where they have vowed privacy and then shared user data with advertisers, researchers, and other users of the service.
- **Does the company have a business interest in compromising your privacy?** What is the company's business model? If the company's business would be aided by it compromising your privacy, it's worth a hard look at whether it is a company with whom you want to trust your images. What do I mean here? Advertising. Is the company's income dependent on advertising sales? If so, they probably have an interest in the content of your photos, as the nature of your photos helps them build a more complete profile of who you are. The more detailed the profile of you as an individual, the more valuable that data is to sell to advertisers who will pay higher prices to better target their advertising. If they make money through advertising, be skeptical about claims of privacy.

Privacy is an important concern for the companies you choose to use for your photography ventures beyond cloud services, and we'll revisit additional privacy topics later in the book.

…and Security

A company's security goes hand in hand with its claims (and record) around the privacy of your photography. After all, a strong privacy policy is meaningless if the security measures in place aren't adequate to keep your private data, well, private.

What does security mean in the age of cloud computing? One major misconception around computer security is that the bad guys are interested in your data. That's generally not the case. Alyssa's senior pictures or Ben and Jamie's wedding photos aren't of much use to a computer hacker… your data is worth far more to you than it is to them. Instead, the hackers want to take control of your accounts to use the storage or computing power for nefarious purposes. It could be some kids screwing around or it could be someone with more evil intentions. Some common-sense precautions will help keep your data safe.

In general, keep a couple of things in mind:

- Use different passwords for each site and service. You can use a password manager such as 1Password or LastPass to make this easier.
- Use strong passwords. While "Lucky69" might be a great name for your band, it's not a great password. Again, a password manager can help suggest and manage good passwords.

Password security is the major thing that you'll need to worry about as someone using cloud services, but there are a host of other security concerns for cloud computing. Thankfully most of those are only concerns for the companies such as Amazon and Google that are operating cloud services and don't require any work (or advanced knowledge) on the part of you, the photographer, as their customer.

Computers and the Creative Realm

What about the creative factors in photography? Can we automate creativity with a computer? Although some interesting projects are being done with computer-generated artwork (we'll explore some later in the book), when it comes to photography I would say that we can't automate the creative factor directly… but there's a big caveat to explore.

While we might not be able to automate creativity, we *can* automate other processes in our camera or with software that lead to effects that previously were a result of creative vision. Let's look at a couple of examples, which we'll explore again later as we talk in more detail about specific technology:

- A light field camera can capture a scene and then adjust the focal plane *after* the capture, which allows one to change which part of the image is in focus. This was previously a creative decision that had to be made at the time of capture.
- Modern smartphones such as the iPhone can capture an image in Portrait mode and then apply a traditional lighting pattern effect to the image (via software) after it's been captured. Perhaps you're not sure which sort of lighting would be most flattering for your subject? Instead of having to make this creative and somewhat subjective decision ahead of time, you don't have to worry about it as you shoot and you can compare options later.

Consider the possibilities of being able to edit the creative aspects of an image hours, days, or even years after the shoot. One interesting exercise for current photographers is to use the latest editing and image processing software to work with older photos, often leading to new and improved results. We can do this now with the basic technical edits in our post-processing software, but imagine being able to adjust creative aspects as well.

When we look at the issue of using computers for a creative effect, you won't be surprised that if we gather a broad range of photographers together, we'll find a broad range of opinions on whether this is problematic. I generally find such arguments about how computers are ruining photography boil down to folks who either hate change in general or hate the fact that something that was once hard for them personally might now be easier for someone who is just coming into the field.

Instead of lamenting that we once put a lot of energy into learning something that used to be difficult, we can celebrate the new opportunities and efficiencies presented by our modernized photography world.

An Evolution

Our modern photography world is evolving, and the evolution pulls together technology, creativity, industry forces, and our definitions of what is photography itself. We've seen a variety of forces change the photographic world over the years. Digital cameras, the internet, smartphones, home printers, and software advances have all shifted the course of photography. As photography has changed, photographers choose how they embrace the new developments. Most changes make the world of imaging accessible to increasing numbers of people.

As artificial intelligence changes photography, longtime photographers will see their worlds change. New people will enter the industry. New companies will pop up, and others will fade away, or exit the imaging-related portions of their business. In some cases, companies that weren't previously considered to be photography companies will end up in that category because of technological advances placing them squarely in the photography world.

As we explore how artificial intelligence impacts the various dynamics of photography, keep in mind that a traditional view of "what is photography" might not encompass some of our new computer-driven reality.

3. The usual moniker for cameras with sensors providing the same size image capture as with a 35mm film camera

3. Words: What Do They Even Mean?

There are a handful of terms that often get thrown around as we dive in to explore the latest technological advancements at the intersection of photography and computers. If we're going to talk about AI as it relates to photographers, we need to learn the lingo. Don't worry, I'm not going to try to explain advanced computer science… this book is intended to be an easily accessible read for the average photographer.

That said… let's take a look at some AI-related terms and explain what they mean.

Artificial Intelligence (AI)

I've encountered a variety of definitions of AI, but I like this one from Builtin.com:

> *Artificial intelligence (AI) is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence.*

Things that "typically require human intelligence" are generally situations where one reacts to unknown or potentially changing external factors. What sort of questions might we ask when it comes to the "intelligence" of computers as it relates to photography? Consider these:

- Who is in this picture?
- Where was this photo taken?
- What's the general mood of this image?
- I have two similar photos… which looks better?
- What's in the image?

We'll explore these questions and others in the coming chapters.

As we look at the evolution of computers (both hardware and software) over time, one will note that things that were once considered to be artificial intelligence are no longer in that category as the things that they do are no longer considered to "require human intelligence." One example is that of optical character recognition (OCR)… if you scan or photograph a paper containing text, a computer can analyze the text and identify the characters and words. Once considered an advanced computational function, it is no longer remarkable and any smartphone has apps available for scanning and identification of words in various forms.

The field of artificial intelligence is broken down into three broad categories.

- **Artificial Narrow Intelligence** (ANI), also known as weak AI, is AI that's focused on a single, narrow task. This is the only type of AI that exists today, and this is the AI that we'll be exploring throughout this book. You've worked with narrow AI applications already, whether it's Siri interpreting your voice command or a photo application that detects faces (and perhaps identifying which images are of your spouse Rosa versus those of your uncle Eduardo). ANI can handle tasks as well as a human for an extremely limited set of tasks.
- **Artificial General Intelligence** (AGI) is a type of AI that is about as capable as a human for a general set of tasks. The model used to create AGI is that of the human brain, but we don't yet have a good picture of exactly how the human brain works. When we start looking at AI that's about as good as a human brain, we get into murky waters. Humans make a lot of mistakes. Some of them are harmless, but some of them cause serious injury or death to others. If the AI is "as good" as a human, can we expect similar results? Do we accept that, since humans make plenty of errors? Big questions here… which thankfully are beyond the scope of this book.
- **Artificial Super Intelligence** (ASI) is a futuristic AI that, in theory, could far surpass the ability of the human brain to make intelligent decisions. What differentiates humans from machines now is our ability to make complex decisions such as those involving emotional relationships or creativity. Why stop with computers that are *as* smart as we are if we could make them *smarter*? This is a while off into the future, but it's something to consider.

Computational Photography / Illumination / Imaging

We increasingly hear the term *computational photography* tossed around, whether it's on podcasts or in an Apple keynote presentation touting their latest iPhone camera upgrade. We know these words separately... we know it has something to do with computers and photography. Together the term computational photography refers to image capture and processing using digital techniques instead of optical ones.

If you used early digital cameras before the optics got very good, you may have used the digital zoom to extend the reach of your camera's lens. This was technically computational photography. You may also remember that digital zoom on the early digital cameras was not great, to put it mildly[4]. Thankfully we now have more modern examples to draw from, and the capabilities have grown along with the technology.

A more recent example of computational photography includes the ability to manipulate an image's lighting after capture, as is possible with the Portrait Lighting features on an iPhone. This is known as *computational illumination*.

Computational imaging refers to the enhancement or creation of an image using indirect means, to create something that wasn't originally present in the original capture. One example (which we'll explore a bit later in more depth) is the ability to generate a 3D representation of an object from a still photo.

Artificial Neural Networks (ANN)

If one starts looking into AI, you're bound to hear the term "neural network" before too long, whether it's a discussion of computation or a reference to the neural processing unit built into the last few generations of iPhones.

To understand a computer version of a neural network we first look at the biological neural network in our brains. The human brain is made up of about a hundred billion neurons, which are connected to other neurons through connections called synapses. Brain activity involves large numbers of these neurons communicating via synapses to form our thoughts and perform various activities of intelligence.

An artificial neural network is similarly constructed in that it involves a large number of individual calculation functions which communicate with each other and pass the information along to eventually reach some sort of a decision or output. These artificial neurons receive information from other sources, assign different weights to the pieces of information, combine the information based on the weighted input and then pass that information along to yet more functions. The individual math calculations being performed might not always be complicated on their own, but the combination of vast numbers of these calculations happening and then being synthesized together is what allows for artificial intelligence to perform operations such as identifying whether a photo is of a person, a cat, or a train.

Think of a neural network as thousands or millions of tiny computer decisions all working together to make decisions normally made by a vast number of brain cells.

Machine Learning

Within the field of artificial intelligence, we find *machine learning*, which describes computer algorithms that are designed to "learn" and adapt based on data received. Instead of a software engineer having to write code to analyze a photograph against thousands or millions of specific criteria, the software developer can write code that analyzes in a more general fashion and uses its own rules to learn about an image or another area of interest to photographers. The machine learns on the fly, adapting its code execution and actions taken in response to what it learns as it goes.

Supervised Learning

For many of the photography applications for machine learning, we want the computer to be able to identify an output based on an input. For a specific example, let's imagine we want the computer to analyze a photo and identify types of fruit in the image.

The task of *supervised learning* is the process by which a computer will learn to map inputs to outputs using a

mathematical model. We start with a training data set of the given input. We then identify features (attributes) of that data and identify which output is correct for the given input. Let’s break it down with our fruit identification example.

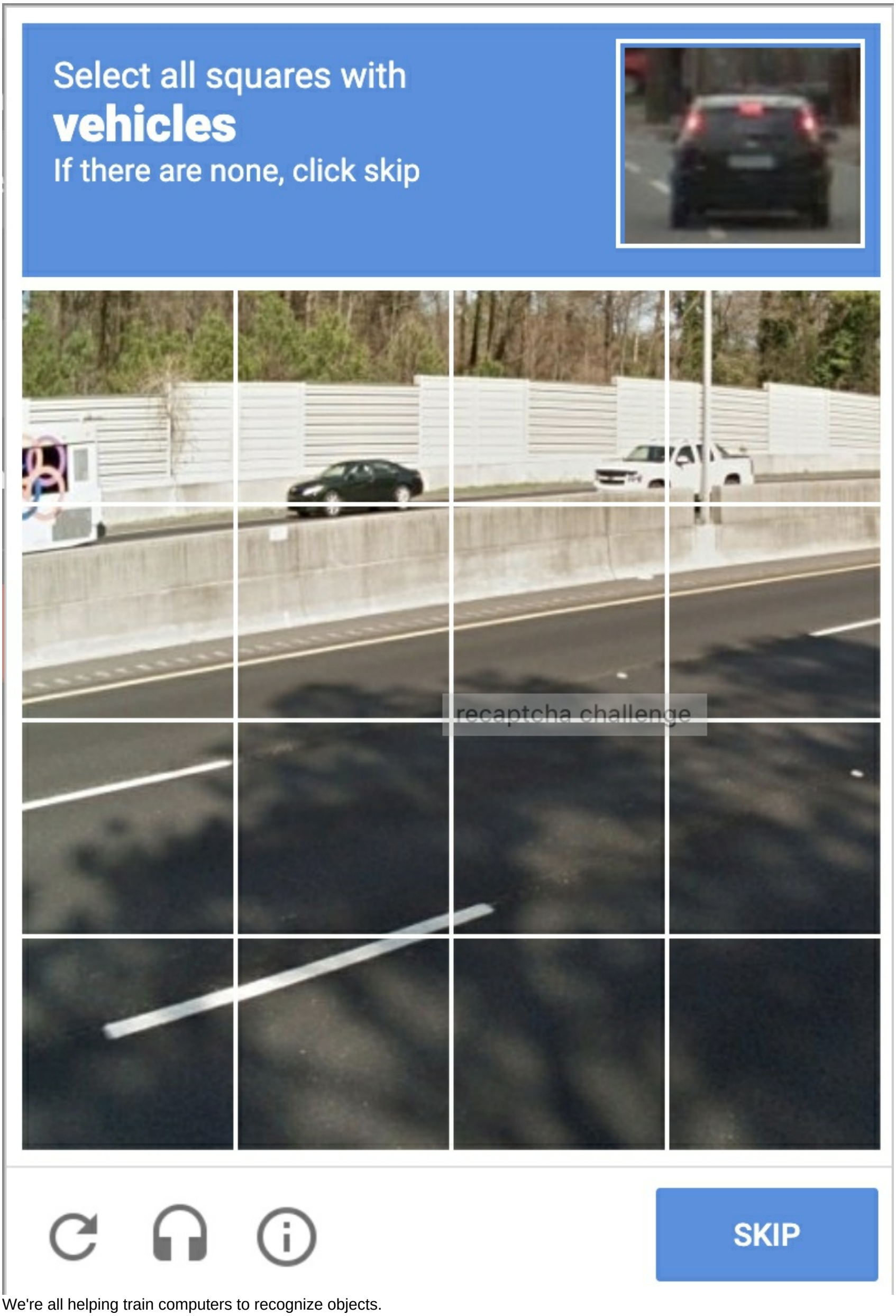
Our **training data** for fruit identification might consist of sets of example images. One set is photographs of apples. Another set contains photographs of bananas. And yet another has photographs of oranges. For our purposes here, this world only contains three types of fruit.

In training the machine, we identify **features** (attributes) of the training images. What might that mean if we're talking about fruit?

- Color values will differ... apples, bananas, and oranges will each contain different RGB color values in varying amounts.
- Shapes will help identify the fruit. Object recognition can learn that apples and oranges are of a similar shape with some subtle differences, whereas bananas are a different shape altogether.
- Different fruit will have a different texture. While an apple is smooth, the peel of an orange has a bumpy texture.

These various features in our training images will help the algorithm decide whether an image contains one of our three types of fruit (or might not contain any of them at all). After we map our training data to its various features, a **learned function** is created using a chosen algorithm that can make decisions based on arbitrary input. We want our function to be able to classify any pictures of the fruit in question, not just those pictures that were part of our training data set. The computer will be able to make that classification based on the various features we’ve identified and how those map to the various outcomes.

This all sounds academic, but let’s make it real with a situation that surely every one of you has experienced. You fill out a form online, perhaps to log into a service, and you see an image similar to this one:



We're all helping train computers to recognize objects.

How does a cloud computing algorithm understand how to recognize objects in an image? It learns from training data sets. In the example shown in the preceding photo, I was asked to help train a computer to better recognize vehicles, with this captcha presented by the Google-owned reCAPTCHA[5] service.

As an interesting aside, we might wonder why Google wants to recognize vehicles. Or crosswalks. Or fire hydrants. Or bridges. Or any number of other things that we're frequently asked to identify in these sort of CAPTCHA

images. Why might a computer need to know about these things? Consider a self-driving car and the sort of environment it will need to navigate on its own. It will need to know about other vehicles, bridges, crosswalks, and similar objects so that it can safely move about the roadway system. By asking large numbers of humans to help the machines learn about various things the self-driving car will need to understand, it greatly accelerates the machine training processes and allows the computers to become “smarter.”

Training data doesn’t only come from things like CAPTCHAs. The sources for machine learning training data, especially in the photography world, are often much broader sets of information and ones that aren’t without controversy. Over the past few years, more than one company has come under fire for scraping photos from various sources. Even when one obtains images to be used in training data in legitimate ways, the data isn’t always the best representation of the situation at hand.

The quality of the machine learning will be influenced by the quality of the training data.

Later on, we’ll take a look at a company using an ambitious set of data for an AI project: they’re looking to use any publicly available photograph they can find. If you think that idea sounds both simultaneously awesome and also a bit frightening, you’re not alone.

The Machine Learned, But How?

It’s one thing to point some algorithms at a figurative pile of photography and tell it to learn and start recognizing things, but for the programmers to truly understand the machine learning it can be helpful to have the AI system explain what it did and why.

Consider the example of identifying a dog in a photograph. If a computer can identify a dog, we can have the algorithm keep track of how it came to that identification. We can ask the computer what parts of the photograph (what pixels) were used to make the identification. We can identify which attributes helped it come to the conclusion. Was it the ears? The tail? The paws? By understanding more about how the decision was reached, researchers and engineers can gain insight into how to refine and train the algorithms. By studying this sort of information, we gain insight into how machines learn.

Sometimes our discoveries with machine learning might surprise us. In a paper presented in May 2019[6], researchers found that while humans primarily use shape to make general identification of objects, computers primarily use texture (not shape) to make identification determinations. Upon further analysis, there are some reasons why this can make sense. There are far more pixels in an image dedicated to variations in texture when compared to the pixels that define the general shape or outline of objects. One can also consider that at some level, texture is just very fine shapes. Considering texture versus shape also can inform the imagery used for training the computers. As photographers, we’re used to digital noise in our images, which is generally thought of as an annoyance. We can lessen the visible noise with control over our ISO setting in our camera or by using noise-reduction software tools. For AI learning purposes, if we know that texture is important for machine learning, we can consider that image noise will be far more disruptive to having accurate texture than it would be for having accurate general shape recognition.

There’s a bit of a circle of learning where humans build algorithms, the computer uses those algorithms for machine learning, and then humans can learn from how the computers performed their work.

4. In general, it was unusable. If you had to resort to using the digital zoom, you were going to end up with a blurry photo.

5. CAPTCHA stands for the Completely Automated Public Turing test to tell Computers and Humans Apart.

6. https://www.quantamagazine.org/where-we-see-shapes-ai-sees-textures-20190701/

4. It Starts in the Camera

In talking about artificial intelligence and how computers will perform many of the functions that once required various skills from photographers, we aren't looking purely at software. Well… technically we are, but we're not just talking about traditional *computer* software or *smartphone* software. We also are seeing growth in the capabilities of the software built into our cameras. AI will make new things possible, and existing things easier. Let's look at some examples of this technology in place today and some examples that are likely to be coming soon.

I Spy… a Face

I'll admit I've been the victim of some camera mistakes in the past. As a professional event photographer, I'm often working with a dynamic situation involving multiple people moving around within my camera frame. As things shift around, occasionally I've been bitten by a focus issue, where the autofocus system ended up focusing on something that wasn't the main subject of my image. It's especially frustrating when some random object is sharp and faces are blurry.

Now we have the technology to help with this problem.

One of the most common examples of AI technology being applied in traditional cameras is with software that performs face detection in images and does so in real-time as part of the camera's viewfinder system. In a capable camera and with face detection enabled, the camera will use the information about faces in the photograph to assist the autofocus system in doing what it can to ensure those faces are in focus and sharp.

We don't think of face detection as being a bleeding-edge feature anymore, but it's possible because of earlier technology gains. As we look at other developments with AI and photography, some of the things that now seem novel or futuristic will likely be thought of as ordinary several years down the road.

Eye Autofocus Tracking

A more recent example of AI-powered assistance in our cameras is the concept of continuous autofocus that can lock onto a subject's eyes. Known as *Eye Autofocus* and *Eye AF*, this brings together a few different camera features for a powerful result. Eye AF is what happens when you blend a continuous autofocus mode along with eye detection. The camera can identify and lock onto a subject's eyes, then continuously adjust autofocus so the eyes remain in focus as the subject moves around the camera frame, including movement toward (or away from) the camera.

Originally this eye detection and focus was designed to help us create more technically correct human portraits, locking focus onto the eye which needs to be sharp. Recent developments now add eye detection that works on animals as well, ensuring that Fido's portraits are as great as possible.

Eye AF is present in mainstream cameras that have been available for a few years from several manufacturers. Sometimes the improvements from artificial intelligence aren't in making futuristic, brand-new things possible, but in easing a pain point or making it easier for photographers to focus on storytelling and composition rather than technical details such as manually managing a focal point.

Your Style, Straight Out of the Camera

All interchangeable-lens digital cameras from the past couple of decades have featured some form of in-camera image styling options. Canon calls them *Picture Styles*, Nikon calls them *Picture Controls*, Olympus users will know them as *Picture Modes*, and Sony folks will reference *Creative Styles*. Regardless of the name by the particular brand, they all boil down to the ability to configure various in-camera settings to be applied to images as they're created.

These settings vary by brand and camera but typically include the ability to adjust things such as contrast, saturation, sharpness, and color hue. Most cameras ship with a variety of preconfigured styles for different artistic

effects. You're also able to go into the camera's settings via the menu system and edit or create your own styles. Want your images a bit more saturated than the defaults? Make it so. Want things to be softer, with less contrast? That's just a setting change.

While I know relatively few people who significantly edit the styles on their camera to change their image capture, any serious photographer is going to spend a bit of time editing their images after they import them to the computer.

Here's where I start musing about future possibilities that could be enabled by machine learning. Imagine if some software (from your camera manufacturer) could examine the edits you've made to the images from a given camera, and then run learning algorithms against those photos as a data set. The output from this learning could then feed back into your camera into a custom, updated-over-time picture style based not on some arbitrary numbers in the camera menu, but instead based on the actual edits you've made to your photos.

In theory, this would mean the photos coming from your camera would get closer and closer to your desired style over time. The amount of post-processing work should decrease since the images will be closer to your finished result as soon as they're created. And as your editing tastes change or your style evolves, the algorithm would continue to incorporate that new data and new direction into your "look" as created by the camera.

Exploring further, this wouldn't have to be limited to a single style. Perhaps you have a look for your landscape images and a different one for your black and white street shooting. It should be possible to teach the camera about the multiple styles such that your editing information was incorporated appropriately to create images reflecting your desired style.

Imagine that… a camera that created images uniquely your own, right in the camera, based on not just your hunches about what you think you want, but based on the actual edits you've made in the past.

You Can Blink if You Want To

How many pictures have been ruined by blinking?

Or by a subject that's not looking at the camera?

Yes, photographers can edit images after the fact to alter eye positions or facial expressions. Photographers who make images of groups are quite familiar with the "head swap" where we manually composite a better look for someone into an image so we have the best possible group shot. Portrait photographers know that the eyes can provide a key connection point with the subject; when the subject of an image is looking into the camera lens, the image viewer often feels a connection as well. Eyes that aren't looking at the camera lens can make a subject feel distant or disinterested.

Apple knows this and has addressed a subtle issue with its FaceTime video call application. When on a call, folks typically look at the middle of the phone or tablet screen, but the camera is off to one edge. This shift in focus, although it's only a few inches, is noticeable to the viewer on the other end of the call. With iOS 13, Apple introduced a feature called "attention correction" which uses AI to adjust the image of the eyes in real-time so that the person viewing a FaceTime call sees the person looking directly at the camera.

While Apple's attention correction feature works with video images during a FaceTime call, another company has been working on opening one's eyes during still photos, so to speak. Facebook released a paper in 2018 discussing the technical aspects of such an effort, where they use AI to "create" open eyes on images where the subject was blinking. A technique called a *general adversarial network* (GAN) samples from other images of the same person to "paint" the eyes using technology. Because Facebook often has a large collection of images of a known subject from which they can learn, this technique works well for their purposes.

Although it's companies such as Apple and Facebook currently experimenting and building these features, there's no reason why this AI-driven technology won't be coming soon to our cameras and their built-in software. If you shoot a burst of five images of a group, why should it be up to you to swap out eyes when someone blinks? Canon, or Nikon, or Panasonic's in-camera software should be able to handle that automatically, meaning when you review your images you'll only see the corrected versions.

Less manual work correcting blinking eyes means more time doing more interesting things, whether that's working with your clients, hanging out with your kids, or playing video games.

Hands-Free Shutters?

We talk about artificial intelligence and how it's going to help us as we capture an image, or in image selection, or editing… but what about an AI-powered camera that decides when to open the shutter in the first place? Instead of a photographer pressing the button with their finger or a remote trigger, the camera surveys the scene and decides when is the right time to trip the shutter and capture the image?

Is it possible?

It's already been done. In a consumer product, nonetheless.

In 2018 Google released the *Google Clips* camera, a $250 device that (among other things) is cleverly named. It has a clip on the back of the camera and captures clips of your life. But of more interest to us, it uses artificial intelligence and machine learning to decide when to capture images.

The concept is simple and yet requires a fair amount of technology to pull it off. Once you've turned on the Clips camera, it "watches" a scene and when it decides there's something interesting going on, it'll capture a 7-second video. The video is a 15fps set of stills, so you can pull out any given image as a still photo. This sounds great in theory, but how well does it work?

Initial reviews on the device from major tech and gadget blogs had rather mediocre results. While the concept is fascinating, it turns out the AI in the device wasn't so great at identifying when something interesting was happening. While Google is fairly tight-lipped on exactly how the device decides what is photo-worthy, the reviewers did a fair amount of experimentation. The device can be trained using a Google Photos library, with the theory being that if it recognizes familiar people, it will use that information to capture images. One reviewer noted that it seemed to capture someone smiling more than when they weren't. Another wanted to see if it could take photos of her dog, but it turns out it took more photos of the dog's rear end than its face. One found moderate success with images of their rabbit.

Google Clips wasn't quite ready to be a useful device for the masses (it was quietly discontinued in late 2019), but it represents a direction for the future that can't be ignored. AI to recognize important subjects, scenes, or activities represents interesting possibilities for photography. We can imagine various scenarios that would evolve beyond acting like a slightly intelligent security monitoring camera.

I expect we will see more developments in the future around technology-based methods to decide when and what should be captured.

No, Not the Bad Kind of HDR

High Dynamic Range (HDR) photography became widely known in the early 2000s after a couple of software manufacturers released their wares to make it easily accessible. HDR allows photographers to merge images in post-processing to create images with a higher dynamic range (that is, more gradations in tone between pure white and pure black) than could be captured by a single frame of the cameras at the time.

A typical HDR pattern would be to combine three images, with one image exposed "properly," one exposed a couple stops brighter, and one exposed a couple stops darker. The software would then merge the images, using the variations in exposure to keep more highlight and shadow detail than would otherwise be possible with a single frame.

The technology was interesting… it could give images a "pop" that wasn't there before. However as it gained in popularity and folks realized they could create surreal technicolor hellscapes with one click of their mouse, many photographers turned all of their images into surreal technicolor hellscapes. It was a garish look that gave HDR a bad name, but not all HDR is bad HDR. Many photographers use it sparingly and with discretion as they strive for realistic images. HDR allows them to bring back additional highlight and shadow detail that would otherwise be lost. This information is the detail that is visible to our eyes but not to our sensors.

The use of HDR for manually creating images that more realistically matched what our eyes can see was a predecessor to more recent feature developments.

Both Google (with Android) and Apple (with iOS) have spent many years working an HDR mode into their camera

functionality for their mobile operating systems. For a while, this was an optional feature, but in recent years the feature has become so good that it's now enabled by default, and a user has to make an explicit choice to create images *without* the default HDR setting. You might be making HDR images on your smartphone without even realizing it. When you capture what appears to be one image (with a single tap of the shutter button) the camera is actually capturing multiple images and using various AI techniques to best align and match the color to create the final photo that you see. These HDR images aren't garish art pieces, but rather are using the HDR technology to create better "everyday" images that more closely match what your eye can see.

Seeing in the Dark

In late 2018, Google unveiled a new feature for Android smartphone cameras on its Pixel phones, adding a mode called Night Sight [7]. The feature allows one to "see in the dark" using this camera mode, and it's a significant leap forward over the traditional methods used to obtain brighter pictures in the dark. Consider that a nighttime scene often contains more subtle variations in tone than the same scene does during daytime; instead of asking what is light and what is dark, we ask what is dark and what is darker. The challenge with small sensors such as those found on smartphones is that there is a notable amount of image noise as the captured image is converted and rendered as a photograph. This image noise appears as little specks or grains of brightness variation. Too much of it and the image suffers.

The easiest way to overcome this noise is to capture an image with a longer duration (shutter speed), but a longer exposure introduces problems with motion. The most common problem is one of camera shake or movement, a problem solved by many photographers by using a tripod. But tripod use is decidedly *not* the normal use case for smartphone cameras, and even for users of interchangeable-lens cameras, tripods are often used for a minority of the images captured. Another source of motion will be the movement of the subject(s) of the photo. Occasionally we want to see this motion blur (for example, capturing an image of a race car streaking past) but generally we don't want our subjects to be blurry. Blurry people… blurry trees… blurry concerts… blur is often the enemy. Avoiding this blur means we can't simply make brighter images by always using a longer exposure.

When I surveyed my tech- and photo-savvy friends in the Android world for a one-line summary of Night Sight, they said "HDR on Steroids." It's a gross oversimplification, but it gets at the technologies involved. Of course, to create nighttime images that look as realistic as daytime images, with the tiny sensor found in a smartphone, requires far more smarts than just basic HDR. Let's consider the computing involved, and how this is far more complex than just simple image merges. The AI starts its work as images are captured.

Intelligent Variable Shutter Lengths

We've established that:

- Long exposures give us more light, but also have the undesired side effect of motion blur.
- Short exposures help ensure we have a sharp image but don't let in much light (especially at night).

In building the computations behind Night Sight, AI begins analyzing the scene when the camera app is opened, even before someone presses the shutter button. While analyzing the scene, the AI looks at the image being captured. Does it detect motion in the frame? If so, the computer understands that to reduce motion, it'll want shorter shutter durations to minimize any motion blur. On the other hand, if the captured frames don't show evidence of motion blurring, then the AI understands that longer exposures can be used to better capture available light.

These changes in exposure length happen automatically… the app chooses optimal shutter durations to best create photo frames to be merged, much like an HDR image created manually from three manually chosen exposures. Merging images is nothing new, but the smart merges in modern apps can intelligently pick the "right" part of the various frames to combine together in more advanced ways than were possible with the software even a few years ago.

Fixing the Yellow, er White Balance

Night Sight isn't purely about a better HDR. One of the other challenges we often face with nighttime images is representing colors accurately, especially in environments with artificial or mixed light. Moonlight on a beach is pretty straightforward and doesn't provide much of a challenge for any camera today, but if we start adding

manmade lighting sources to the scene, whether they're halogen, tungsten, or fluorescent, the story gets more complicated and we often end up with photographs with odd color casts that don't reflect what's seen by our eyes.

Why is this so hard for a camera? Part of the reason is that our human eyes are very good at something called *color constancy* which allows us to see a color as being consistent, regardless of the lighting situation in which we view it. If we view our blue car, we see the car as being the same color regardless of whether it's in the white snow at noon or at the beach in the golden rays of sunset. Unfortunately, when a camera captures a scene and we then view a photograph, our eyes will often see a tinted world, with an overwhelming color cast based on the type of illumination present at the time of capture. To overcome this scenario, a camera adjusts the color to compensate for the color cast based on the type of lighting. This is known as white balancing and allows our eye to view an image as if it were in a neutral lighting scenario.

Traditional white balance algorithms in cameras often have issues in low light, especially when mixed with non-natural lighting such as the sodium vapor lights that are pervasive in our cities at night. Google developed a new learning-based algorithm for automatic white balance that is used to handle these special cases at night.

AI has made the Night Sight mode possible both in learning the appropriate shutter speeds to capture to create the best-possible combined image for a given scene, as well as in how to ensure that color is as realistic as possible. If we take out the AI aspects, the nighttime smartphone images wouldn't look nearly as great.

And iPhone Too

In 2019 Apple added its own features for nighttime and low-light photography with a new Night Mode as part of its native camera app on the iPhone. Whereas Google frequently blogs openly about the technical aspects of their AI features, Apple remains secretive, but we can look at what's offered and make some conclusions about what's happening behind the scenes.

The Android Night Sight features are invoked explicitly by the user when they choose their camera mode. For Apple's Night Mode, it is applied automatically by the iOS camera app when it detects a low-light situation such as dawn, dusk, nighttime, or a dark indoor room. In the camera app, users see a crescent moon icon to the upper left of the viewfinder. As with many things in iPhone photography, Night Mode works by combining features of the image sensor, the image processor chip, and software using artificial intelligence to end up with a pleasing, realistic image. Whereas Android's Night Sight is occasionally criticized for making images look unrealistic (a nighttime photo often ends up brightened such that it no longer looks like nighttime), Apple's Night Mode photos tend to end up in a more realistic state, closer to what our eye can see.

When Night Mode is in effect, the camera is capturing multiple images of varying shutter speeds, up to and including a long exposure in order to capture as much light as possible. As we've mentioned, a long exposure invites the possibility of camera movement. This is taken into account when the image is processed, and the iPhone can use data from the device's motion sensors to determine just how much camera shake was present during the shot. After the various images are captured, the software takes them together, identifies which are the sharp and usable portions of the various frames, and then processes them into one finished image. When all goes as planned, the finished image has realistic colors, sharpness, and a minimal amount of noise.

The success of nighttime photography (smartphone or not) often depends on the ability to stabilize the camera and make a long exposure that allows for the available light to be recorded on the sensor. The iPhone's optical image stabilization comes into play, and Night Mode will identify how long the camera should be stable to capture the ideal image. This duration is displayed to the user near the Night Mode indicator, showing the number of seconds the camera should be held stable. For a handheld shot, this might be a second or a few, but if the iPhone's sensors detect that it's on a tripod or otherwise in a stabilized situation, it can capture for much longer (up to 30 seconds).

As we look at Apple's Night Mode and Android's Night Sight together, they both address what has been traditionally one of the weaknesses of smartphone (and other small-sensor) photography. Low light scenarios, such as those at night or in dark rooms indoors, have always been an area where larger sensors were the only hardware capable of making great images. That is no longer the case. Innovations in hardware, including the camera sensors and image signal processors, have made the smaller-sensor captures better. But the real power has come from software innovations, with the camera software able to better identify "good" and "bad" portions of captured images and to combine multiple quick-succession captures to create one resulting image that can rival those of devices with bigger hardware sensors.

When Digital Zoom Got Good

One thing digital photographers have learned over the past twenty years is that optical zoom is fantastic. With moving lens parts, you can get a closer view of your subject, and the resulting image should be of the same quality regardless of how much you zoom. Digital zoom, on the other hand, hasn't been so great. Whereas optical zoom changes the field of view of what makes it to the camera's sensor, digital zoom has traditionally "zoomed" the image by cropping to a smaller area of the sensor and then enlarging the resulting image to fill the frame. As as we've all seen, when you enlarge a photo, the computer doesn't always do a great job of interpolating. Digitally-zoomed images often suffer from pixelation and appear less sharp than optically-zoomed images. As the image is expanded, the software attempts to fill in the missing pixels, but it has done so in a way that lacks the detail found in the original image or the actual scene.

With the Google Pixel 3 camera introduced in late 2018, Google introduced a "Super Res Zoom" feature that redefined digital zoom and resulted in high-quality images through the use of machine learning to improve the software used to create the resulting zoomed images.

Previous digital zoom images resulted from taking a single frame, cropping, and then attempting to enhance the photo through software. With the Pixel 3, when a user pinches to perform a digital zoom, instead of using a single source, the camera uses multiple source images (from a burst capture) to perform the interpolation. We discussed earlier in this chapter how HDR and multiple-image capture is becoming the norm, and the Super Res Zoom feature is another case where multiple source images provide plenty of raw material for software to use to create images that aren't possible from a single frame.

The concept of using multiple images to create a higher-resolution or zoomed image isn't new with the Pixel 3, but rather this is the first time we saw it applied in everyday photography use. The ability to use multiple similar images to zoom has been a technique in astrophotography for a while. That process is known as "Drizzle" and was originally developed for the Hubble Deep Field camera observations made by the Hubble Space Telescope. Since then, it's been applied for other astrophotography uses, and a Google search reveals everything from specialized software to YouTube videos explaining the concept. You may have heard of cameras that use sensor shifting to capture slightly offset images to combine them to create an image with a higher resolution than is captured by the sensor in a single frame.

The software used for the Super Res Zoom feature mixes these things together. The slight movements in a smartphone during a burst capture mean there will be source images that are close, but not quite identical, allowing for the creation of the higher resolution image to use for the zoom. While image blending has been possible for a while, we are just now starting to see it as a feature in mainstream capture devices (smartphones and other new cameras). The software is finally good enough that it can combine the images quickly, apply solid noise reduction, and use various AI and machine learning enhancements to produce a solid final picture.

The Dynamically-Updating Camera

When you bought an early digital camera, what you bought was what you got. Whatever software features (and bugs) were present was what you had to work with. At some point, digital camera manufacturers added the ability for photographers to apply firmware updates. Firmware refers to the software-on-the-hardware. It's the software that runs the camera, processes images, enables the computerized autofocus system to work, and otherwise controls the various other electronic features of the digital camera. Firmware updates usually occur by downloading an update from the manufacturer's website and installing it to the camera (either via a USB cable connection or by putting the update onto a memory card and using an in-camera update feature).

Until around 2015 or so, firmware updates were usually used for one purpose: fixing bugs. A firmware update would typically resolve some software issues. Sometimes these would be user interface issues, or perhaps there would be a tweak to make a slight improvement to the autofocus for some particular scenarios. These bug fixes were valuable to folks who might have been experiencing a specific problem, but many photographers happily owned cameras for many years without ever applying any updates.

Some of the mirrorless camera manufacturers started using firmware updates to not only resolve problems but to introduce entirely new features. This has now become common across several brands of cameras. A firmware update might bring new (and often valuable) enhancements to the camera's capabilities. Here are some examples of features that didn't exist in cameras as they originally shipped, but were added via (free) firmware updates:

- A face selection feature in the Fuji X-T3
- Real-time Eye AF for the Sony Alpha a9
- Focus bracketing for the Fuji X-H1
- Tethered shooting capability for the Panasonic GH5
- External flash configuration in the camera on the Sony a7R III
- Focus stacking for the Olympus E-M1

There are many other examples from various manufacturers, but with the increased software capabilities in our cameras, we’re fortunate in that we can often improve our capabilities just by installing a firmware update. We’ve seen that many of these updates are related to features powered by artificial intelligence.

7. Night Sight has since been added to other high-end Android phones

5. Showdown: Traditional Cameras vs. Smartphones

Some thoughts on traditional cameras versus smartphones are in order. Many in the photography industry have seen this as an us-vs-them battle where those darn kids with their newfangled iPictures are coming in and screwing up "real" photography as practiced by old-timers with big cameras.

It's not an us-vs-them situation.

Thankfully attitudes seem to be shifting, at least into the realm where folks realize that both interchangeable lens cameras and smartphones can play valuable roles in the photography world. At a hardware level, they both have optics such as lenses and sensors. And they both have processing chips — image signal processors — that perform a variety of calculations on photographic data. These processors aren’t new, in fact, you probably have seen some of them advertised with their branded names on feature lists for your favorite camera model:

- Canon has *Digic*
- Leica has *MAESTRO*
- Nikon has *Expeed*
- Sony has *Bionz*

Recent smartphones also have image signal processors. For both traditional camera and smartphone image capture, these processors help with things such as sharpening, noise reduction, and demosaicing of the images as they are saved to memory.

Attitudes and hardware aside, I see some big themes on the software side of things when it comes to the adoption of new technology — including artificial intelligence — when looking at the world of standalone cameras (DSLRs, mirrorless, and so on) and the world of smartphones.

Traditional Camera Manufacturers Suck at Software

If we look at the big players in the traditional camera market such as Canon and Nikon, they don't have a great track record when it comes to the software part of the digital photography equation. How often have you thought “Wow! This DSLR menu is so intuitive and easy to use!” I’m guessing… just about never.

Software is becoming more important to the image capture process to take advantage of AI possibilities, and camera manufacturers’ ability to compete could become constrained if their software can’t keep up with their hardware. The manufacturers of DSLR and mirrorless interchangeable-lens cameras currently have a big optical hardware advantage over smartphones; their bigger sensors and huge lenses enable high-resolution, precise image capture. Physics is on their side.

We know what’s possible today using the software and interfaces of these cameras. What could be possible in the future if they can bring cutting-edge software into the picture (pun intended)?

Smartphone Makers Are Innovating in Software

With traditional camera vendors having a big advantage when it comes to the hardware of sensor size and lenses, smartphone makers have (by necessity) turned to software to help them make great images. With a disadvantage when it comes to physics, smartphones (currently) have an advantage when it comes to software.

Smartphone camera software isn't better because of chance; it's better because it *has* to be for the images to be able to be used for meaningful photography work.

Both traditional camera manufacturers and smartphone makers will continue to work on hardware enhancements for their products, but optical hardware innovations run into limits due to physics in a more definitive way than software innovators will run into the limits of what’s possible with artificial intelligence and the software.

Stack Your Chips

One other area of disparity between traditional camera manufacturers and high-end smartphone manufacturers such as Apple and Google is evident when we start looking at the processing power found in modern smartphones. Gone are the days where the only function of a digital camera's "computer" core was to record the data from the image sensor onto the memory storage of the device. There is now serious processing power in your smartphone, and a lot of that power is there for computational photography.

The term *CPU* (central processing unit) has been around for decades and is widely known as the brains of a computer for processing calculations and instructions. Modern smartphones also contain two other major processing components which help with photography and image processing. The first is the previously-mentioned image signal processor, and the second is a *neural processing unit* or NPU. The NPU is a chip specifically focused on artificial intelligence tasks. Traditional CPUs can perform AI calculations, but doing so puts a strain on their resources, leading to slower operations using a lot of power. On a smartphone, that power usage means battery drain, so a traditional CPU isn't well-suited for the AI needs of smartphone photography.

A neural processing unit in a smartphone can overcome the limitations of traditional CPUs, and allow for advanced AI calculations to happen on the mobile device without excessive battery drain. While many mobile services depend on sending data over the internet for cloud processing, in other cases handling the computations on-device presents advantages for users. One advantage is in the areas of privacy and security; if the data never leaves the smartphone, it eliminates any possibility for interception or misuse in the cloud. Another advantage of handheld processing could be that of speed; for many tasks, the on-device calculation will be faster than depending on a network connection.

CPU speed is still important, but for many mobile tasks, including photography, a solid NPU is just as helpful.

Everyone Will Look to AI

Regardless of current strengths and weaknesses, anyone involved in the camera industry is looking to artificial intelligence and other software as ways to improve their future products. Canon, Nikon, Sony, and other interchangeable-lens camera manufacturers haven't *had* to innovate as much in software at this point of the market's evolution, but they can't get complacent and assume it won't be needed in the future. You'll see that these companies will be hiring software engineers with expertise in AI to help them enhance their products. Will we see AI-focused chips for neural processing make their way into traditional cameras? In mid-2020, Sony announced the first image sensor with AI processing as part of the sensor itself. This sensor is launching first in the industrial market and it's unknown when this technology will make its way to the sort of cameras used by you and me.

You'll also see continued emphasis on software developers with AI skills going to work for companies such as Apple, Samsung, and Google. AI and software will continue to be important as they compete not just with the bigger cameras, but with each other. With photography being one of the most important uses for smartphones, market success depends on the ability to have great images come from their cameras.

What remains to be seen is how the mix of traditional vs. smartphone cameras plays out, and how that impacts the future of where software and AI advances reach the market. Current trends would indicate that AI and neural processing innovation will continue to come from the smartphone world, while traditional camera manufacturers will continue along a path of incremental improvements to their decades-old technology and with software playing a minor part in their story.

6. Meanwhile, Outside the Camera…

Thus far we've reviewed quite a bit of material around the state of artificial intelligence as it relates to camera hardware and image capture. Whether we're talking traditional dedicated cameras or newer smartphones, we see that AI is making big advancements and changes in how we capture images.

Cameras aren't the only types of photography hardware that incorporate artificial intelligence as part of their design, however. Let's take a quick look at two other pieces of imaging hardware.

Follow the Bouncing AI

Photographers who use a speedlight flash on their cameras know that if you point the light directly at your subject, you end up with harsh lighting, problems with red eyes, and often a "deer in the headlights" look to the image. If one angles the flash head upward toward the ceiling (or to the side toward a wall), a photographer can "bounce" the light and obtain a broader, softer light that helps flatter the subject and generally results in a more pleasing image.

Until recently, choosing how to adjust a flash to create a bounced light was pretty much a matter of guesswork and experience. In 2018, Canon brought some modernity to that skill by introducing a speedlight (model 470EX-AI) that uses AI features to help photographers obtain a good bounce angle. Using what Canon calls "AI Bounce", the flash makes measurements of the distances between the flash head and the surrounding ceiling and/or wall being used for bounce, and automatically adjusts itself to provide the best angle.

Will auto-adjusting bounce angle on a speedlight revolutionize photography on its own? No. But like many of the improvements coming as a result of AI being integrated into our digital photography world, it will eliminate a friction point and make it a bit easier for photographers to find success with the technical aspects of creating solid images.

Add AI on Your Hotshoe

While the big traditional camera manufacturers such as Canon and Nikon haven't been touting very many new AI features in their flagship DSLRs, there's at least one product on the market attempting to add some AI goodness to your traditional DSLR. The Arsenal unit[8] mounts on a hot shoe and plugs into a camera from companies such as Canon, Nikon, Fuji, or Sony.

Arsenal sits on top of your camera, evaluates the scene, and interfaces with the settings on your camera to best match the situation at hand. Arsenal applies object recognition and uses that information in tandem with known information about how to photograph certain subjects. As an example, for a fast-moving subject such as a bird, it will automatically ensure a very fast shutter speed.

In addition to object recognition, it can also make your camera's HDR functionality even smarter, automatically making adjustments to ensure that highlights aren't clipped or shadows don't end up blacked out.

In short, Arsenal brings many of the AI functions we know from modern smartphone camera software into use with more traditional camera units. This gives you the hardware power of bigger lenses with some of the software smarts of AI-driven camera systems.

8. http://witharsenal.com

7. “Better” According to Whom?

We might assume that photographers are always hoping to create better photos. And we'd be right. I can't think of a time when I woke up, grabbed my camera, and went out to make pictures and hoped they were of lower quality than previous images I've made.

What is Better, Anyway?

If we set technology aside for a moment, we realize that "better" is a subjective measure.

Judgments about the quality of a photograph are often dependent on any number of factors including technical elements, the viewer's familiarity with the subject of the image, and the viewer's relationship to the photographer. The relationship factor is one often overlooked by casual photographers or those just starting to get serious about their images. It turns out the reason our images might get a ton of likes and comments on Facebook from our family and friends isn't so much that they're great images, but rather that those folks are family and friends and they want to make us feel good.

As we grow in our work, we often seek external validation beyond those who are socially obligated to be supportive. Perhaps we enter a photo contest. It could be an online contest that gathers anonymous votes or an offline situation such as a county fair. Either way, we put our work into the pool of entries to be ranked against the other photographs, and we know there will only be one first-place winner.

Is the work that places first, second, or third in a photo contest "better" than the work which didn't? Hopefully. It must've been better in the eyes of the judges, based on whatever criteria were being used for evaluation. Are the judges qualified? Sometimes they are, or sometimes the general public might perform the judging role. Is it a verdict on which art is better?

Does the definition of "better" get any clearer if we look at a more formal and (in theory) more objective judgment of photographic quality? If we take the next step beyond a photo contest and instead look at a formally judged competition, with trained judges, does the evaluation of a photograph's quality become less cloudy? Take, for example, the International Photographic Competition hosted by the Professional Photographers of America (PPA) every year. The PPA judges all have years of training and experience, and instead of ranking images against each other, they're scoring work based on what PPA has defined as the twelve elements of a merit-worthy image[9]. Factors such as composition, technical excellence, impact, and storytelling work together to help the PPA judges determine a score (up to 100) for an image as it stands alone, rather than being compared against other work directly.

Although the work isn't compared directly, an image scoring 95 should be "better" than an image scoring 75, shouldn't it?

You might be wondering why, in the middle of a book about artificial intelligence and computers and technology, I've spent several paragraphs asking questions about how we evaluate what "better" means in the context of photography. I ask these questions because as we look at artificial intelligence being the use of computers to perform tasks once only thought possible by humans, we pose a key question:

Can a computer judge "Good" photography versus "Bad" photography?

Can it make those determinations on an objective scale? What about judging two images against each other? Can it use those determinations for a productive purpose?

Neural Image Assessment

To cut to the chase, yes, computers can evaluate "better" as understood by humans at this point. Let's explore that concept, and why we care.

Some of the factors that make an image "better" than another image are technical concerns where it follows logically that a computer could make a technical evaluation. Things like image sharpness (or blurriness), excess

noise in a photo, or compression artifacts are relatively easy for a computer to identify and use as factors in considering whether or not an image may be seen as being of good quality.

Beyond the technical aspects, artificial intelligence can also help us understand and evaluate what's most pleasing to the human eye from an aesthetic standpoint. Researchers at Google released a paper at the end of 2017[10] where they explain how they successfully trained AI to make those judgments, previously thought to be subjective and only possible by human eyes. This neural image assessment algorithm was able to accurately rate and judge image quality in forms that came very close to the ranking made by humans. The study used images that had been rated by over 200 individuals, allowing for an evaluation of the rankings and the ability to plot the ranking distribution for comparison between man and machine.

These computer-driven rankings were found successful in two different forms. Firstly, when comparing similar images against one another, such as when a base image was compared to versions of the same image where enhancements (or detractions) had been made. The second form of successful AI ranking was in taking a data set of similarly categorized images (such as landscapes) and ranking them in order from most to least visually pleasing.

The computer can tell which of several photos of a given landmark is more attractive, as well as being able to know if that landmark photo is more or less appealing than images of other landmarks.

The Applications

Why do we care whether a computer can evaluate whether an image is "better" according to humans? Do we spend enough time looking for "better" photography such that artificial intelligence can play a big role in making our day-to-day lives easier?

Better Inside the Camera

Some of the uses for which we want a computer to be able to identify the better image are for use with image capture and processing. Let's look at a couple of simple examples from the iPhone world:

- Apple's Live Photos mode captures short video and identifies which frame of the video is the best image to be used for a still photo.
- Burst mode works much like a continuous shutter release on a traditional camera, capturing a series of images, continuously. When reviewing images in the Photos app, these images are stacked together, with iOS displaying what it feels was the best image of the group.

These two examples refer to situations where the AI is going to identify the best image and then save the chosen frame for later human use. If such identification can be made quickly, AI can use that judgment and feedback to make real-time decisions about image capture. Comparing bracketed images almost instantly, the camera could make adjustments to the exposure to ensure the most pleasing image is captured.

In a different use case, Square (the company best known for its easy mobile payment processing using a smartphone or tablet) entered the photography market in mid-2019. Businesses can send products to be photographed with the Square Photo Studio service, which provides product photography at a low rate made possible in part by automation and robotic control of the camera. Eliminating the human photographer and the expenses of traditional commercial photography lowers the fee paid when compared with traditional product photography rates. The resulting images won’t be high in creative value but should meet the needs of businesses who need typical (read: not too creative) catalog-style photos. At launch, Square indicated that while the camera capture process was automated, staff would review the images and make the selections of which photos get sent back to the clients. As AI image quality judgment continues to improve, this seems like a situation where it could be employed to evaluate the series of captured photos and choose the “best” of the bunch.

As we move deeper into our discussions of AI, keep this ability in mind: when we’re talking about cameras that can capture numerous images at a fraction of a second, the ability to determine image quality — and which image is the “best” of a bunch — can be key in our cameras and their software being able to quickly capture and store the best image possible as we make pictures.

Image Selection

Whether it was a weekend of landscape work, a senior portrait session, or an evening event photography gig, one of the first things a photographer often does after a photo session is to cull through the collection and identify the images to be processed or analyzed further.

Depending on the number of source images, this can be a time-consuming process. I know from personal experience that when photographing a conference or trade show, I'll have several hundred images (or more) from which to make selections… instead of spending hours reviewing images, what if AI could help me?

It can. There are now multiple software programs on the market such as AfterShoot and Optyx that use AI to perform various evaluations and identify the best (and worst) images of a batch. It will be interesting to see if these features eventually make it into the built-in feature set for popular image management programs such as Adobe Lightroom or whether they remain as specialized third-party utilities.

Image Management

What about *after* the immediate post-processing and when we're managing our image collections in the future? There are opportunities here as well for photo management applications to employ AI technologies to help identify and surface our best photos.

Adobe Lightroom is the most widespread image management software in use among serious photographers, offering a wide range of photo management, editing, and publishing abilities. Thus far, image management is mostly a manual affair in Lightroom. It relies on the photographer to sort images, search for them based on metadata, or configure pre-saved collections (essentially saved searches) based on criteria known in advance.

Other than Lightroom, other popular software used for image management includes the native Photos library on iOS and Google Photos on Android. In addition to allowing users to browse photos by date, both systems now offer some sort of intelligence in helping to surface photos that may be of interest to folks. They both have "Memories" features, and Apple Photos has a "For You" section that attempts to bring forward the most interesting stuff.

How often do we capture photos and they get stuck in a digital library somewhere, never to be seen again? By using AI — ideally based on preferences learned from our individual behavior — these apps can help us enjoy the images that could mean the most.

9. https://www.ppa.com/events/photo-competitions/the-12-elements-of-a-merit-image

10. NIMA: Neural Image Assessment - https://arxiv.org/abs/1709.05424

8. After Capture: AI in Post

We've reviewed some of the interesting things that can be done by artificial intelligence as part of the image capture and selection process. Let's turn our attention to additional applications for the technology once we've moved onto the post-processing and editing phases of our photography workflows.

Many of the uses for artificial intelligence in post-processing and other photo editing comes from the ability to recognize shapes, objects, faces, and other things in our photographs.

Select This

Folks who have worked with Photoshop (or any serious image editing application) know the process of selecting items can be a skillset all on its own with a non-trivial learning curve. Numerous tutorials, book chapters, and YouTube videos have been created to help folks learn how to select items in Photoshop, and then how to adjust and refine those selections because very rarely are they correct on the first try. The selection tools have evolved over the years, and now one of the first challenges when trying to select something in Photoshop is selecting which of the many selection tools to use…

In the past, selecting an object (let's say a dog sitting on the ground, or a hat on someone's head) would involve using one of the selection tools, drawing a rough path around the edge of the object, and then adjusting and refining the selection until it was "close enough" for your purposes. In late 2019 Adobe announced another new selection tool for Photoshop, this one powered by artificial intelligence used for object recognition. The new AI-powered selection tool can be used for selecting objects. With the aptly-named Object Selection tool, you simply draw a big selection around the item, it applies AI object recognition, and you usually end up with an accurate selection of the item.

Artificial intelligence is used to identify what object is the target of your selection. This selection ability can be quite powerful because it doesn't apply only to straightforward objects such as a ball or a book, but can also apply to the selection of people or pets. As most of us know from experience, hair is one of the hardest things to select. Human hair… pet hair… any sort of hair is pretty much impossible to select manually and requires using various selection-assistance tools. There have been multiple iterations of improvements in the built-in Photoshop selection tools, and there are third-party plugins such as Topaz Mask AI that are single-purpose tools to assist with selection. Adobe is bringing AI to their tools as well, and photographers will benefit.

As much fun as it's been to spend an hour masking hair to cut out a person from a background, I'd much rather do it in a couple of clicks with a tool that's smart enough to identify hair versus background texture.

HDR for Better or Worse

We already discussed HDR a bit when discussing in-camera AI capabilities, since, at this point DSLRs, mirrorless, and smartphone cameras all have built-in HDR functions (some of which are enabled by default). But if we're looking at how AI is used for image editing and post-processing, we ought to mention HDR here as well.

HDR was one of the first widespread software applications where the computer was able to do something automatically that previously required a *lot* of manual work in Photoshop. HDR didn't invent the concept, but it made it widely accessible with just a few clicks. Before HDR, a similar process would involve making the bracketed exposures, bring them into multiple layers in Photoshop, the performing a complicated blend to make the tones from each layer based on how they work together for a finished image with the maker's vision for highlight and shadow detail.

When the HDR software application merges images and performs tone mapping, it's making mathematical calculations based on the tones in the images. These calculations decide how much brightness should appear in a given part of an image, with the goal being that the resulting image better represents what we can see with our eyes. As the person using the software, we have the ability to control various aspects of this process, and we get to influence the outcome. The goal might not be for a realistic image, but rather for an artistic interpretation of the scene… an interpretation that could vary greatly from reality.

Early versions of automated tone mapping often led to images with some telltale undesirable characteristics. The overall image would often be flat and lacking in contrast and one might find a “glow” around the edges of objects. AI advancements have been able to minimize these distracting artifacts that are a byproduct of the tone mapping process by analyzing the image and identifying things that didn’t look natural.

The result is that tone mapping software in 2021 produces (by default) images that are more realistic and are less likely to have the stereotypical HDR look than the same software from five or ten years ago.

Out of the Shadows...

While HDR in general looks at mapping various exposures to create a result more closely resembling the dynamic range we can see with our eyes, there's one specific case of altering the light and tone in a portion of an image with practical applications for casual, day-to-day photography.

I'm talking about using software to remove unflattering shadows. Sometimes these shadows can ruin an otherwise good image. Hats with brims, nearby foliage, or other objects can cast harsh shadows across faces in casual portraits. A project by Google, Google Research, and the University of California - Berkeley has identified an algorithm that can help repair these shadows and create more pleasant lighting on subjects' faces[11].

By using AI, the software can restore lighting on the face of the subject, resulting in a usable and pleasing image. These shadowy pictures are often the result of casual snapshots, and this is a great application where AI can help the everyday person making images with their smartphone or other consumer cameras.

...And Into the Sky?

Artificial intelligence applies across photographic genres. While many of the examples discussed so far are focused (forgive the pun) on portraiture, there are also AI applications for landscape photographers. One of the most common situations is in performing sky replacement. Every landscape or outdoor fine art photographer has stories of images with interesting foregrounds and dull or distracting skies. As someone who lives in the Pacific Northwest, I can't tell you how many otherwise-interesting images I have with flat gray skies. It’s a large number.

Sky replacement using traditional photo editing software has traditionally been a two-part process: masking and merging the sky, and then performing adjustments to the image so the sky and the foreground look good together as if there were captured as one scene.

AI-powered editing software such as Luminar and Photoshop can make this a breeze. The previously-mentioned masking capabilities help make the sky swap straightforward, and AI can look at the overall lighting of a scene and make adjustments to ensure that the resulting image looks realistic. These capabilities have gotten quite good; Luminar can even alter foreground reflections to match changes made with AI sky replacement.

Content-Aware Magic

One example of AI in post-processing that's fairly widely known is the content-aware fill feature available in editing software. Broadly popularized by Photoshop, there are similar features in other applications as well. Content-aware fill allows a photographer to "fill" a portion of an image with computer-generated imagery that (in theory) should neatly blend in and match the rest of the image in a way that makes it look natural and isn't obvious to the viewer that it wasn't part of the original photo. Content-aware fill is often applied in situations such as:

- after the removal of a distracting object from an image, it can be used to fill in the area where the object was removed.
- extending a sky or background to create "more" image than was present in the original capture.
- a panoramic or stitched image where there are gaps at the edge of the photo that must be filled to result in a clean rectangular image.

Content-aware fill (from various vendors) works similarly in that there can be two sources of information for the

application to know what to do to fill in the missing pieces of an image. The more straightforward source is in the software “looking” at the sections of the image around the area to be filled and using similar colors and textures to fill in the missing gaps. This was the first technique used for early content-aware fill applications and works pretty well for things like skies or extending a printed photographic backdrop.

The second potential source for content-aware fill to use is for AI to identify what's in the rest of the photo, apply some machine learning to understand what is likely missing and use that information to generate the missing portion of the image. This is a newer source for content-aware image editing, and as you might imagine is the area with the biggest opportunity for improvement and growth. Much like other machine-learned photo editing applications, advancements are being made frequently in this technology. As I write this chapter, Adobe just announced another round of content-aware improvements being released into their editing software both for still and video images.

Treating People Differently

A bit further in the book, we'll dive deeper into AI facial recognition (and what it means for the future of photographic metadata) but this ability is also useful at a broader level that isn’t tied to a specific person. Being able to recognize a face means that AI can recognize the greater concept of a person.

Consider how you might edit a photograph containing people along with other elements. The various adjustments we might make to an image such as color, contrast, clarity, or textures are the sort of adjustments where we wouldn't want them blindly applied across an entire scene containing faces along with other elements. One might want to smooth the skin but leave the background alone. Or one might want to add some punchiness to the areas around the subject, while not mucking with the skin tones or texture. With traditional image editing software, this situation would require creating a mask to isolate the person from the background. After creating a mask, the edits could be made without affecting the other part of the photograph.

If AI can easily isolate the person, the mask becomes automatic, making these sorts of adjustments much easier and much faster.

The next level comes when software identifies not just what is a face, but what are various parts of faces, and knows how to adjust them appropriately and selectively. This software already exists… there are plugins designed for portraits that can recognize various aspects of a face and apply editing techniques automatically. With a single click, the software can:

- identify the face in an image
- make overall skin-smoothing adjustments (but only to the skin)
- add some crispness to the eyes by making adjustments to the sharpness and saturation (again, only to the eyes in the image)
- go beyond the overall skin-smoothing and apply a bit more wrinkle reduction to specific areas

This is all possible because the software can look at a face, identify various parts such as the eyes and mouth, and then make conditional adjustments to certain areas automatically. Like most such automatic editing features, it’s not perfect, but if the software can get us 80% or 90% done it certainly makes big inroads into the amount of work needed to apply selective editing.

You’re probably catching on to a common theme to many of the AI-powered features for photo editing: understanding various parts of an image and the subjects contained therein, and then making decisions and automatically performing editing tasks on those parts of the image.

Improving the Scan

Although almost every image is made digitally now, there are still countless older photographic prints which are being digitized for use or preservation. This involves scanning the photograph to make the conversion from analog paper to a digital file we can edit. Even on photographs of good quality where restoration isn't needed, the scanning process is one where we introduce potential image degradation.

You might not be surprised to learn that artificial intelligence can help.

Google has developed technology in an app called PhotoScan (available for both Android and iOS) that uses AI to fix a couple of the most common problems introduced when scanning photographs.

Consider glare on an image. It's pretty common, especially given the glossy nature of most photographic prints. Even photo finishes that aren't considered glossy (such as a lustre or matte finish) will reflect lighting in our environment. In a photograph that will be scanned, those light reflections will show up as glare. The general method that PhotoScan uses to remove glare from an image is that if it captures multiple photographs (by moving the camera) of an image, the glare will move around as the camera moves. As the glare's position moves around the image, different sections of the photo will be obscured by the glare in the different captures. After capturing several frames with glare in different places, the images can be merged, with the software identifying which areas have glare in each image and using the non-glare areas from other images to repair and composite until there is one finished image without any of the glare defects.

The second issue addressed is when a photograph that has been scanned isn't perfectly flat. It's common for a printed image to have a bit of a warp or bend in the paper, and this can introduce a bit of distortion into the image. Some of these distortions might be subtle, but we're not just fixing them for the sake of fixing them. The computational photography work to fully deal with the glare requires that the image be flat so that various sections can be properly aligned. Google uses a technique called optical flow -- a computer vision representation for motion -- to correct areas affected by the non-flatness of the image. As these areas are addressed, it allows the glare-removal and repair methods to work effectively.

Whereas glare removal once required complicated image editing techniques, it can now be performed on mobile devices using an AI-powered app.

Restoration and Colorization

Having just discussed the process of scanning older photographs, let's also explore AI technology being used in the process of photo restoration. Photo restoration is used to "fix" defects in older images that have been scanned. These images might have significant issues including scuff marks, tears, pencil or pen marks, smudges, holes, or other problems. Traditionally, this work has been done by a photo editor using manual processes to identify the image defects and then use various techniques (including content-aware fill mentioned above) to restore the image as it would've appeared undamaged. A photo restorer might choose to colorize the image, applying color to various areas of the work to make it appear as the scene would have in real life as opposed to the black-and-white or sepia tones which were a limitation of the film available at the time.

Let's look at the process and data available from one AI-based restoration project: Computer Vision by Mail.ru. They've shared their information on their process, which contains three high-level steps[12].

The first part of their AI-based photo restoration project identified defects in the image. These could be wrinkles, tears, scuffs, or other blemishes on the photograph. If a photographer were manually restoring this image, they'd start with the same step… identifying the defects. From these identified defects, the software creates a mask (the same sort of mask we might create in Photoshop or another image editor) defining which areas of the source image are "good" and which areas of the source image are "bad" and need to be restored.

The next step of the restoration process is to correct the masked "bad" areas of the image. A human retoucher would use various tools in Photoshop (or a similar application) to correct the image. Depending on the size and nature of the defects, they may use a clone tool, healing brush, or hand-drawn brush strokes to restore the damaged or missing parts of the photograph. For an AI-based application to perform similar tasks, it will rely heavily on content-aware techniques such as those described previously. Using the context of surrounding areas of the photograph, it will fill in or enhance the defective areas of the photo, restoring it to a blemish-free state that should look similar to how the original image would've looked many years ago. One interesting aspect of the AI-based solution is that because a mask exists to know which areas of the photograph are "good" and which are "bad", the infill process can only sample from known-to-be-good areas, resulting in a more correct restoration. This aligns with how the process would be done manually; a human photo retoucher wouldn't use the bad areas as a source for retouching.

The final component of the AI-based process is one of colorization. This is typically done by analyzing a black-and-

white image and predicting the red, green, and blue (RGB) channels, but in the Computer Vision example, they chose to write their own algorithm for better results. Their plan worked well, and they leveraged additional AI to refine their colorization. They used a neural network that could evaluate an image and determine if it looked realistic. Based on the evaluation of the colorization effort, adjustments could be made until they had a more realistic image than would've been possible without the use of artificial intelligence.

The AI-driven restoration and colorization of images represent yet another area where an AI solution will be able to do a substantial amount of work in an automated fashion. Like many of the AI applications discussed in this book, we aren't yet at a point where it's going to replace humans, but rather will be able to automate some of the more basic tasks, allowing retouchers to focus on the finishing touches or to work on scenarios where the AI isn't yet able to solve the problem.

Groups Are Made of People

We've talked a fair amount in the book about how artificial intelligence can merge images in useful ways, whether the merge occurs at capture or afterward.

If we talk about areas of photography that have been impacted by the smartphone, we have to talk about a popular type of casual photography that barely existed beforehand. That's right… with or without a stick.

I'm talking about the selfie.

Do you know what selfies are really good for? Taking a picture showing you in some environment. And do you know where selfies often miss the mark? When you want to take a group photo. Sure, you might be able to get one or two other folks in the frame, but that's about it.

While patents don't always lead directly to products, in late 2018 Apple applied for a patent on technology that would allow photos to be merged into a selfie. The result would be that you could take a selfie with someone who wasn't there, and software would add them into the photo in a realistic fashion.

This came to light as we were a couple of months into the COVID-19 lockdown in my area. We could have socially distanced selfies in isolation. Brought to you by software that can put photos together on its own.

Don't Go Chasing Watermarks

If one wants to get some photographers into a spirited debate, get them started discussing whether a visible watermark is useful on their work. Watermark advocates will claim the mark helps prevent copyright infringement, and those who are opposed will note the watermark is distracting from the subject of the work.

Like many other aspects of photography, the watermark debate gets more complicated as we introduce artificial intelligence into the picture. Let's look at how AI and watermarks interact in two distinct situations.

Watermark? What watermark?

Recall all that content-aware technology we talked about previously that can be used to generate and patch and restore photographs? The same technology can be applied to "restore" a photograph to its pre-watermark state as if the mark never existed. Examples showing automated watermark removal are widespread at this point. With large collections of images bearing the same watermark (as is done with stock photography image collections), those images can be fed through machine learning to easily identify the watermark pattern and remove it automatically.

Watermarks were always about keeping honest people honest because there was a fair amount of work involved in removing them. With the ability for content-aware AI tech to remove them quickly, typical watermarks are even less effective than ever. If someone wants to swipe your photo and it has a typical watermark, AI will make it easy for them to do so.

…Or We Make Them Better

While traditional watermarks are easily defeated using AI to eliminate the watermark text and replace it with appropriate fill material, there are also innovations in making watermarks harder to detect and defeat. One such method, published by Google in mid-2019, involves adding a subtle and random warp to the watermark text for each image[13].

With each image's watermark being unique, it's not simple for machine learning to identify the repeated pattern, meaning that it can't simply bulk-fill as it could with a consistent watermark across a series of images. Instead, the computer must look at each image and deduce which areas are affected by the watermark, then compute how to eliminate and fill that area in a method that attempts to look natural and remove any trace of the watermark removal. This process must be repeated for every image, which makes it far more machine-intensive than when there's a standard watermark in play. The challenge here isn't just one of computing power… after all, computing power is cheap anymore. The challenge is that without having a large data set from which to train the computer and identify the watermark, the computer will be far less effective at identifying which areas are part of the watermark and which areas are part of the image.

The result is a warped watermark that is much harder to be removed via artificial intelligence. Attempts to do so result in images with noticeable visual blemishes, making it clear that manipulation has attempted to remove a visible watermark.

This is unlikely to be a permanent solution to machine-proof watermarks but provides an adequate safeguard for the time being. As technology continues to advance, AI will likely be able to defeat and remove such warped visual watermarks, and creators may be forced to look elsewhere for solutions that provide a visual indication of copyright.

Whither Traditional Textures?

One long-standing way that many photographers add an artistic take on their images is with the use of textures added when editing the image. Whether it's added to the entire photo or only a selected area, a texture can help set the mood for a photograph. Soft flowing textures can accentuate a portrait of a sleeping newborn, for example. Hard, edgy textures can add some grit and grunge to an urban scene.

Traditionally, photographers obtain textures in one of a few ways:

- They photograph them, which generally only happens as the opportunity arises. One can't always go find the desired texture organically.
- They use free textures available through a variety of sources. Although this provides a wider selection than shooting the textures themselves, they're still limited by what's offered.
- They purchase textures. This is the most common situation for photographers who use a lot of textures, as the commercial sellers of texture files make it relatively easy to browse their offerings and find a suitable texture to match your desired artistic outcome.

With most textures, what makes them interesting is that they feature similar (but not identical) patterns repeated throughout the image, and blending them with a photograph results in a consistent mood but doesn't give the look of something automated or computer-generated.

Can artificial intelligence be used to assist or replace traditional sources for textures? It can, and at least one product already exists for this purpose. In late 2019 retoucher Pratik Naik released a tool called Infinite Texture Panel, which works as a Photoshop plugin to make textures available using AI-powered tools to find the perfect one for an image[14].

Infinite Texture Panel has a few interesting features powered by AI. The first is a feature where it lets you paint onto the image with a general idea of the sort of texture or area you'd like to affect, and it will generate a similar texture and apply it to the painted area. There's also a feature where you can bring in a sample of a commercial texture you don't have the rights to use, and it will find a closely matching texture in its library which you can use instead. It's an interesting feature that might raise some ethical questions, but it's a good example of the power of AI to analyze an image and use that analysis for comparison.

A "randomness" slider in the plugin allows you to decide just how uniform (or not) you'd like the texture to be for your use. You can go for something consistent, or something wacky, or somewhere in between.

This seems like a stepping stone on a path where we'll eventually see AI-generated textures via an easy interface accessible to average Photoshop users. As of writing, experiments are being done in this area, but we haven't yet seen this ability built natively into commercial products.

From Still to 3D

We've talked a fair amount about still photography, and have mentioned video a few times, but the world of 3D imaging and animation is another area that will be pushed forward by artificial intelligence technology. While these applications of AI could likely fill an entirely separate book, there's at least one application where it crosses paths with more traditional photography and bears mentioning.

Researchers at the University of Washington, in conjunction with Facebook, demonstrated in 2018 the ability to create an animated 3D rendering of a person based on a still photograph. The animation relies on several pieces of computation to come together.

First, the computer must be able to identify the human subject in the image. As we've discussed previously, recognition of objects is an area where many advances are being made with AI, and thus picking out the human subject is task one. After isolating the subject, the real work begins, which is mapping the still image into a 3D animation rendering, taking into account various properties of the human body and what's known about movement.

The research puts it all together to animate the subject walking toward the viewer using a HoloLens augmented reality headset. The technical details are less interesting to us than the fact that this is now possible, at least in a research environment[15]. Like everything else we've been discussing, this technology will only get better.

Whether it's working with old images to scan and restore them, or working with futuristic 3D generation, artificial intelligence will assist us across a wide spectrum of photo editing and manipulation tasks.

11. https://www.dailycal.org/2020/08/30/uc-berkeley-researchers-develop-artificial-intelligence-to-remove-shadows-from-photos/

12. https://9may.mail.ru/?lang=en

13. https://ai.googleblog.com/2017/08/making-visible-watermarks-more-effective.html

14. https://infinite-tools.com/infinite-textures/

15. https://arxiv.org/pdf/1812.02246.pdf

9. Alexa, Make a Picture?

Voice recognition and voice commands are everywhere. Some studies indicate that as of 2020, over a quarter of United States households have a voice assistant device (not counting the voice-operated functions built into smartphones). With all the major smartphone platforms now having voice services, it makes voice-driven technology accessible to almost any photographer. Each year, smartphone and computer software companies add more features to their voice assistant services including Apple's Siri, Microsoft's Cortana, Amazon's Alexa, and the Google Assistant.

Voice on Your Smartphone

There are currently a few ways in which you can use your smartphone's voice assistant for photography activities.

Siri can currently open the camera app to a few configurations, with phrases such as:

- "Take a selfie" (starts forward-facing camera)
- "Take a picture" (starts the rear-facing camera)
- "Take a square picture" or "Take a panoramic picture" (starts the rear-facing camera in those modes)

The Google Assistant can fire the camera's shutter directly, allowing you to make images using only your voice. This can be quite handy as a remote shutter control. By default, it will count down from three seconds before capturing the image, although you can also specify the number of seconds in your voice command.

The voice assistants get more powerful when integrated with third-party applications offering voice support. While Siri can't snap a photo, when you use the Halide application[16], you now have this capability on an iPhone.

Seek With Your Voice

While using voice for capture can be of some benefit, voice can also be a powerful way to take advantage of the previously-mentioned search capabilities of our digital image collections.

We previously explored the notion of various AI-powered searches, so let's take those to the next level by invoking them via voice. Many of these can be done today with existing voice services.

- "Show me photos from Christmas 2018"
- "Show me photos of my wife"
- "Show me photos of California"
- "Show me photos of sports"

The search and voice features are even more powerful when you combine various searches into a single command that returns the specified results.

- "Show me photos of Sara in Disneyland"
- "Show me photos of taxis in New York"
- "Show me last year's selfies"

Quickly Show Off Your Images at Home

Have a Chromecast attached to your TV and using Google Home? You can use your voice to tell it to show a slideshow of particular images. "OK Google, show me photos of sunsets on my television." Want to see photos from an album of last year's vacation? You can order that up with a voice command.

The voice recognition is great, but ultimately it's powered by the underlying (AI-powered) recognition and categorization of your images.

While it's not purely voice-related, as we're on the subject of photo slideshows in a smart home environment, we can consider what might be possible beyond explicit voice control. In the 2001 movie *Antitrust*, a futuristic high-tech home featured art display panels on the walls which automatically changed as various folks entered or left a given room. The smart home knew the art preferences of individuals and would change the wall art to their preferred style as they entered a room. With biometric sensors and smart processing, such a situation that was considered futuristic twenty years ago is now entirely possible.

Voice to Camera?

If we're looking at how voice control technology might intersect with photography, we ought to look at the various pieces of photography gear we currently use which might gain voice support in the future.

Consider a portrait photographer in a studio environment. For most folks in this environment in 2021, they'll have various lights set up around their subject, triggered by a remote mechanism either native to the camera or in a third-party system such as the popular triggers from Pocket Wizards. A best-case scenario for making adjustments involves the photographer changing settings at the camera (which might be reflected on the remote triggers), or a more clunky system might involve a photographer needing to adjust each light by walking to it and pushing some buttons.

Now imagine a scenario where we're still using remote-controlled lighting gear, but instead of having to push buttons, we can control those lights by voice. Imagine voice commands such as:

- "Increase main light by one stop"
- "Decrease hair light by a half stop"
- "Increase fill light by two stops"

Instead of pushing buttons on a remote attached to your camera or in your pocket, you could simply verbally give commands to your lights. Joe McNally has affectionately referred to assistants as Voice Activated Lightstands (VALs)… but in the future, we may have voice-activated lighting without needing the assistant.

And why would we stop with lighting? We already have remote camera controls in the form of triggers and intervalometers, why would we not have voice control built into our DSLR or mirrorless camera body? Instead of holding down buttons and turning dials (or even worse, navigating an often-cumbersome camera menu system) why couldn't we just speak to our cameras:

- "Increase the exposure by one stop"
- "Turn on HDR bracketing"
- "Shoot a time-lapse with one image every five seconds for five minutes."

The possibilities are vast.

16. https://halide.cam

But What About…

AI can do some cool stuff. But what about some other concerns…?

10. Meta About Metadata

In looking at how computers can analyze a photo and provide us with information, there are any number of items related to metadata that are of interest. What once was done manually by a photographer, assistant, or digital tech can now be done automatically by a machine. Let's explore a bit about how artificial intelligence, machine learning, and other aspects of the modern computing world will impact the types of metadata we use for our photos.

Image Capture Metadata

As we talk about metadata, it can span a wide range of information about the equipment, situation, and subjects of our photos. Artificial intelligence will affect each of these areas of metadata management.

Pre-Digital Metadata

Before the invention and widespread adoption of digital cameras, film photographers kept various bits of image capture metadata about their images. Some of this metadata is now automated or obsolete.

One key piece of metadata we don't even consider anymore is that of the capture media. In analog photography days, the vendor and brand of the film played a significant role in the look of the images captured. Even beyond the photography world, we saw this creep into pop culture.

> *They give us those nice bright colors*
> *They give us the greens of summers*

Those lyrics were Paul Simon singing about Kodachrome[17] in 1973. Kodachrome had a look. As did other products such as Fuji Velvia, noted for its saturation and fine grain. Knowing the capture media was a notable piece of the story of a photograph.

In the digital world, such a piece of information is irrelevant… our final image isn't affected by whether it was originally recorded on a memory card from SanDisk, or Lexar, or any other manufacturer of digital memory media.

Capturing the Capture Settings

There were several other bits of gear-related metadata that serious analog photographers would record manually.

Beyond the film (manufacturer, brand, and speed/ASA), we would capture bits of information including:

- lens
- shutter speed
- aperture
- lighting equipment used

We would typically record the date and time of the photo as well, which was of particular interest for nature images where the time of day can be a key factor in the lighting situation.

Together this information helped form a basis for image review after the film had been developed. How did a photo look when compared with another and considering the settings? Was something underexposed? Overexposed? Which f-stop gave the desired depth of field?

As we moved into a digital world, we enjoyed the benefits of having much of this metadata captured automatically. The EXIF metadata standard provided the ability for cameras to record this same information (date, time, camera used, as well as settings such as shutter speed and aperture) embedded into the resulting digital image files. What once required a photographer to carry a notebook and manually make note of camera settings was now captured automatically by the digital camera and recorded into the same file containing the picture itself.

Here's an example of image capture data as recorded by one of my cameras and displayed in Adobe Lightroom:

Metadata as shown in Lightroom

Obtaining this metadata required zero effort on my part as the photographer. It’s handled automatically with features built into every digital camera and smartphone. We’ve come a long way from the days of carrying a notebook and having to remember to write down camera settings in the field.

Metadata in the Frame: Keywords

Another important use of metadata is to capture information about what’s happening in a particular image. Unlike information about camera gear, (most) of this sort of metadata isn't yet being captured in any automated fashion. Traditionally this is where serious photographers would use keywords on their images, either stored in a separate catalog program[18] or directly embedded into the image files.

Here are some common examples of the sort of things that might be stored as keywords for various genres of photography:

- Weddings: bride and groom names, venue name, the role of the subject (bride, groomsman, etc), the part of the ceremony, the name of a vendor in the image
- Events: name of the event, venue, city, involved organizations or companies, names of people photographed
- Portraits: names of individuals, style of the shoot, attributes of individuals (woman, blonde, sitting, etc)
- Landscape / Nature: names of locations, names of natural features, type of features (forest, waterfall, insect, etc), time of day (sunrise, sunset)
- Sports: player names, event names, team names, the name of the action happening (bunting, passing, celebrating), locations, stadium names

Can technical advances including artificial intelligence help modernize the world of metadata? They definitely can.

Let’s consider object recognition of what’s in the frame. Apple, Google, and other companies have already tackled this. The technology is available today through products like Apple Photos and Google Photos. Even without tagging your images, you can give add your images into these software programs and then search to find images that match your criteria. Want to see your waterfall images? Turns out they can identify waterfalls. Want to see pictures of sunsets? The same sort of thing… no metadata is needed because the algorithms can recognize those scenes in the pictures and make them available via search.

What about people? Through facial recognition features, it’s now possible to identify individuals to image management applications (such as Adobe Lightroom, Apple Photos, and Google Photos) and the software will then recognize and identify other photos containing those same people.

How does this affect the future of metadata in a world where our computers and phones can perform image analysis? It will mean a world where photographers can reduce the amount of time spent adding metadata to images because much of it will be redundant.

There’s no need to add a keyword to photographs containing Aunt Maria if your image management software can automatically find photos of Aunt Maria. There’s no reason to add a keyword to images containing a waterfall if your software can find photos of waterfalls. Remember all the time we used to spend adding routine keywords to our photos? We’ll get that time back for other more interesting activities.

Where the Heck Are We?

They’ve said it’s all about location, location, location… so how do we track the location of our photos?

If we go back to the turn of the century, it was pretty much all manual, after the fact, and handled via metadata tags on the images. After returning home from a shoot, we’d use our software of choice to add location data to the images by choosing a location from a map or typing in the name of the venue.

If we fast-forward to the present day, there’s decidedly less manual entry of locations due to the proliferation of GPS in the camera world. Most of the major camera manufacturers have GPS available for their interchangeable-lens cameras, whether it is built into the camera as a feature or available via an external device. Our smartphones have had GPS for a decade, and thus any image coming from our mobile devices is going to come off the device with geotagging already embedded[19].

If you have a camera *without* a GPS device, but you have your smartphone with you, some apps can track your position over time and then use that log to later add location data to the images shot with your non-GPS camera. Location data is synchronized using the timestamps of the images to automate the addition of location metadata to the images from your camera.

Let’s look at how location metadata might be impacted by artificial intelligence. We’ve discussed object recognition. What if that object recognition got really good, not just for categories of things in general, but to identify specific landmarks or locations in our photos?

- Your photo of the Eiffel Tower would be automatically tagged not just with appropriate keywords indicating it’s the Eiffel Tower but also geotags indicating the location[20].
- Your picture of a major league baseball game might be recognized and the computer could identify which stadium you were at.
- As you capture an image of the Golden Gate Bridge while standing on Alcatraz Island, even without GPS it could learn where you were.

What once seemed impossible or magic is now possible thanks to computing advances. If we think back to our original definition of artificial intelligence — the computer being able to accomplish tasks once thought only possible by humans — this sort of work falls squarely into this category.

What Just Happened?

Events and circumstances represent another broad area of metadata currently kept as keywords associated with an

image, but with the potential to become automated as AI can recognize situations. Let's break this down based on a few genres of photography.

Wedding Metadata

As one of the most popular and profitable areas of photography, let's see what a computer might be able to figure out after it has been trained with object recognition and a bit of data to help it learn about weddings. As I work through various situations here, it's wise to keep in mind that I'm describing the typical things at a typical wedding and that less traditional events might not fit this pattern. But it's a start, and computers can learn about alternate situations as well.

- **People**: traditional Western wedding attire is fairly predictable with a bride in a white dress and a groom in a tuxedo or suit. A computer that recognizes that pairing in several images, as part of a greater set of images shot on the same day at the same venue, could deduce we have a wedding. And we've just identified the bride and groom. And once we know we have a wedding, what do we have when we have a group of women in nearly identical attire? We have bridesmaids. Similarly, there's a good chance we have groomsmen in matching outfits as well.[21]
- **Cake**: A multi-layer cake or grouping of cakes? Perhaps with figures on top? There's a darned good chance it's a wedding cake. And a photo that involves two people both holding onto an object that intersects the cake? That would be the cake-cutting photo.
- **The Kiss**: Two faces joined at the lips. Surely we can recognize a kiss, right?
- **Flowers**: A computer can easily recognize a bouquet. Put that in the hand of the bride (who was already recognized) and you have a bridal bouquet.

Put these pieces together and you have a decent set of data about the wedding event and its set of images. If you join that metadata with some solid search capabilities, it's not hard to ask your computer or phone to show you particular images or image sets.

Portrait Metadata

While most portraits don't have as much going on as you find in the breadth of images from a wedding, there are still some bits of metadata that can be identified by a computer that has been trained.

We can start with the fact that an image is a portrait… for single-subject portraits, it should be identified that there's a single subject, whether their eyes are engaged with the camera or even looking away. Once we identify the subject in a portrait, we could deduce other aspects about the portrait to better classify the image. Is it a headshot? Are these newborn photos?

Some areas of portraiture will be harder to classify via an algorithm. Are these high school senior photos, or just a portrait of a young woman? And what about groups… are these three women together posed as business partners, or is it a mother with her adult daughters?

While portraits don't drop easily into as many identifiable scenes as the weddings we discussed, modern computers should be able to start making some educated predictions.

Sports and Adventure Metadata

The broad spectrum of sports and adventure photography is another great candidate for AI to do some metadata processing to better identify the activities in the images. As we look at mainstream sports such as baseball, soccer, football, or hockey, there are many attributes to the photos that can be used to sort images into classification by sports. What are those attributes? They're the same things we look at when we view an image with our own eyes.

- What sort of uniform is being worn by the players?
- Are there visible elements that are part of the game? I'm thinking about balls, bats, pucks, or other equipment.
- What about the field? There's a big difference between a white ice hockey surface and a green soccer pitch. And a baseball diamond is certainly a recognizable design.

Beyond traditional sports played on a variety of fields, we can also look at adventure sports. Instead of manually having to add a keyword to an image indicating it involves rock climbing, or scuba diving, or skydiving, we can let the computer do the work. Many of these activities feature scenes that are identifiable to our eyes and can be

identifiable to a computer that's been provided with appropriate training data.

Awards and Ceremonies

While wedding events often follow usual patterns, there might be other opportunities for AI to recognize what's going on in our images of other events such as award presentations or other ceremonies.

When someone's posing holding a plaque or a certificate, we can identify it as being some sort of award or recognition. We all know what a "grip and grin" image looks like, and a computer can figure it out as well.

The primary challenge in metadata for award and ceremony photos is that it's challenging to know just *what* sort of event is involved. Is this an academic award? Is someone being presented with a promotion at work? Is it a photographer being recognized by their local PPA affiliate?

While the computer can often identify the nature of the event, the specifics are probably going to be hard to deduce… at least for now.

What's the Genre?

We've touched on this a bit as I've looked at how computers can classify various sorts of images, but one broad classification can be the photographic genre. While there's no definitive list of genres, the Professional Photographers of America offers thirty-three specialties for which their members can classify themselves in hopes of being found by clients.

Architectural	Buduoir	Child
Commercial	Documentary	Drone
Equestrian	Family	Fashion
Fine Art	Headshot	High School Senior
Landscape	Maternity	Nature
Newborn	Pet	Photo Restoration
Photojournalism	Portrait	Printing
Real Estate	School Portrait	Special Event
Stock	Travel	Tween
Underwater	Videography	Wedding
Wedding Videography	Wildlife	Youth Sports

Photographic Specialities identified by PPA

As you can see some of these are pretty broad and well-known (such as Wedding and Portrait) but others are more specialized. As we look through this list of genres, I can think of common factors within many of these groups that could potentially be identified by AI and used to help classify the photos or assign appropriate metadata.

Meta In a Mood

Another area of metadata that could likely be at least partially assisted by artificial intelligence is the concept of tracking metadata around the mood or feeling of images. If I were to say words like eerie, bright, fresh, cheerful, ominous, or airy, you probably have an idea of what sort of photographs would fit those words. Can we train a machine to make similar classifications? I suspect that we can if we break down the various things that lead to emotion or mood in an image.

One consideration is the overall **tones** of the image… if we consider a histogram, which shows the distribution of tones in a photograph ranging from absolute black to absolute white, what is the profile of a given photograph? Is

this a high-key image? Low key? Somewhere in between?

Are there **faces** in the image? Whether it's a studio portrait, an environmental portrait, an event scene, or any other image with people in the frame, facial analysis and clues can tell us about the mood of the individuals. In aggregate, each one of the folks in the scene can help dictate the overall mood of the image.

Can we train about **natural elements**? Consider the emotional impact of a soft, flowing stream versus whitewater rapids. Both are natural water features, but one can evoke peacefulness while the other emits tension.

The **lines** in an image can factor into the emotions or feelings an image might evoke. Do we have gentle, easy curves, or do we have many hard lines with sharp angles?

While we tend to think of emotions, moods, and feelings as less-defined things in our world, as we break down the elements that evoke emotion in a photograph, we're able to isolate certain attributes. Train the AI on those attributes, train the AI which attributes lead to which emotions, and we've trained the AI on how to recognize things in an image that are very much human, *non*-machine feelings.

Oh… and one other mood consideration…

Are there clowns? Clowns will mess with anyone's mood.

What Can't We See?

While AI-based image recognition and queries can replace many traditional uses of metadata, it isn't going to be perfect and can't handle every scenario. If we depend on artificial intelligence to make determinations based on what it can "see" in our images, it's going to be limited to the sort of metadata that can be seen or inferred from visual elements in our photographs.

Not all metadata can be seen.

Google Photos originally used AI to automatically tag photos with the gender of the subjects in the pictures. These tags were based on a training data set which identified certain visual aspects as associated with a specific gender. In early 2020, Google updated its systems to identify people in images as "person" instead of a gendered name such as "man" or "woman." At the time, Google stated that "we have decided to remove these labels in order to align with the Artificial Intelligence Principles at Google, specifically Principle # 2: Avoid creating or reinforcing unfair bias."

As we look at the possibilities of AI replacing metadata, we should consider that there are times when AI won't be the solution.

All This Metadata: Is It Necessary?

We've taken a pretty substantial look at the metadata that photographers have typically tracked for the images. Very little of it comes "for free" - other than the EXIF date and time information, almost all of the metadata we've discussed was generally either added after the shoot via keywords or other image management techniques, or it required specialized equipment that wasn't standard on most cameras.

Adding metadata to images was a necessary part of a photographer's work so that the metadata could be used to manage the photos. At times we'd need to know the details of a single image, or at times we'd need to be able to search and sort by various bits of data. Consider these typical queries a photographer might ask of their image collection:

- Show me all of my sunset photos
- What images have I made with the mayor in them?
- Do I have any pictures of boats? What about a specific boat?
- What images do I have showing fall-colored leaves?
- Which lens do I use for most of my pictures?

There are a number of questions we can ask with various specifics depending on the situation. We've added

metadata to our images to support asking those questions.

What if we can ask those same questions and get the answers *without* having to spend the time and energy to add metadata? If our image management programs and smartphone apps can provide those answers through artificial intelligence, do many of our metadata practices become things of the past?

That's where we're headed. Before the announcement of Apple's iOS 13, I wrote an article with a wish list for photographers with the new software. Metadata management was near the top of my list, but nothing new was announced in that area — at least not concerning traditional metadata of keyboarding and such. It turns out that's okay. Because instead of focusing on traditional metadata, Apple is devoting its energy to the future. They're beefing up the AI features to make it easier to find and work with images where the computer does the hard work.

There's a quote often attributed[22] to Henry Ford stating:

> *If I had asked people what they wanted, they would have said faster horses.*

Ford was suggesting that true innovation usually doesn't come from asking the customer. When I lobbied for better metadata management in iOS 13, I was that customer. Instead, what Apple is delivering is the ability to better manage our images, but leaving traditional metadata behind. Apple isn't the only one making strides in this area. Adobe is working on similar AI features in their applications, and many other lesser-known companies are doing the same.

Traditional metadata management is gradually going away, with new, AI-based solutions to answer the same underlying questions.

17. https://en.wikipedia.org/wiki/Kodachrome*(song)*

18. Properly referred to as a Digital Asset Management (DAM) application

19. Unless you've disabled this feature for some reason… but if you have, you're probably not concerned much with location metadata.

20. Distinguishing between the one in France vs. the one in Las Vegas would be another challenge.

21. I realize I've just described a heterosexual wedding and that there would need to be additional training and algorithms to also ensure the software was as accurate as possible for same-sex couples and their events.

22. It appears that Ford wasn't actually the source of the quote, but it was cast onto him. https://quoteinvestigator.com/2011/07/28/ford-faster-horse

11. Private AIs are Watching You

One concern with the ever-increasing capability of artificial intelligence as it intersects with photography is the issue of privacy. It’s been mentioned previously, as we talked about cloud services, but it’s worth a broader discussion.

When we discuss privacy, everything is on a spectrum. As individuals, we vary where we fall in our comfort level around *who* should be able to know *what* about us.

At one extreme we find folks who generally don’t care much about privacy and figure their life is on display and that unless we’re talking about something *extremely* personal or intimate, there’s no need to worry about privacy in our everyday lives. On the other end of the spectrum we have individuals who greatly value personal privacy, and feel that companies, governments, and other individuals should only ever learn any personal details about ourselves when there has been an explicit choice to share that information. Most of us fall somewhere in between.

When we're talking specifically about how privacy relates to artificial intelligence, we have tradeoffs to consider. The best machine learning would come from having the most data from photographers and other technology users, including data that might typically be considered private. By respecting user privacy and not using private data for training the AI algorithms, we generally lessen the effectiveness of those algorithms.

What Guides Privacy Expectations?

In looking at our expectations of privacy, we’re guided by two patterns, one of which is more rigid and the other of which can be flexible. Both can evolve. A complete discussion of privacy is a far greater topic than is relevant here, but we’ll look at privacy issues related to photography, and those specifically which might intersect with the role of automation, artificial intelligence, and other technical topics in the evolving photo world.

Legality

Our first privacy pattern are things which are coded into law. These laws vary by country and have evolved over time. Based on my locale and education, my discussions here will be decidedly US- and EU-centric… if you’re outside the Western world you may find different scenarios in your area.

Photography privacy laws regulate when one has an expectation of privacy (and thus not to be photographed) along with what rights someone has concerning the usage of a photograph for which they were an identifiable subject. In the United States, there is generally not a right to photographic privacy when one is in a public space. If you're walking on a sidewalk, or driving your car, or standing in your front yard, it's fair game for someone to snap a photo. Conversely, if you're in an area with a reasonable expectation of privacy, it's considered a crime for someone to try to photograph you. You can’t stick a camera into a public bathroom stall and start snapping away.

If an image contains a recognizable person, there are laws that restrict how that photograph may be used without that person’s consent. Some uses, such as an editorial photo in a newspaper, are permissible even if the subject of the image hasn’t approved the use. In other circumstances, such as using someone’s image in an advertisement that implies endorsement of a given product, one is required to obtain permission from the subject. This is why model releases are obtained from photo subjects with language that explains the terms under which an image may be used .

Most of our laws were written considering photography as being an interactive, human-driven endeavor, with a human brain making decisions about what was happening, when, and to who. Keep this in mind for further exploration.

Social Norms

Beyond the law, our expectations of privacy are also guided by social norms. We consider what sort of behavior we expect from most people in a given situation. You can likely think of several situations that might be permitted under the letter of the law but would be considered creepy or otherwise outside what we expect as normal behavior. One common example is that while photography in public venues is generally allowable by law, someone who shows up at a children’s playground with a telephoto lens and makes pictures of random children is going to raise some skeptical eyebrows as folks wonder if their intentions may be less than honorable.

Laws get updated over time, but social norms change much more quickly. In the early days of the smartphone, it was almost universally considered rude to be using one's phone at a restaurant during a meal, and now as we look around any given restaurant, we're likely to see multiple individuals using their smartphone for photography, often of the meal itself.

In a World of AI and Photography Automation, What Changes?

Until *very* recently, most of our expectations, laws, and norms around photo-related privacy all involved the typical photography model where a human has a camera, is watching the scene, and choosing when to capture an image. The one exception would be the realm of always-on surveillance cameras, but even those were traditionally deployed and used in fairly predictable ways. As surveillance cameras have become more modern, our laws and expectations around them are often tied to the norms of the locale (public surveillance is much more common in the United Kingdom than in the United States, for example).

How will attitudes shift... will they be based on the effectiveness of the technology? When is it good enough? As of mid-2019, the London Metropolitan Police's facial recognition system had a failure rate of over 80 percent![23] Traditional surveillance was the capture-and-review model, where the footage would record to tape or disk, and later be reviewed by a human if there was a cause for concern. The major shift with the increasing power of our camera technology, coupled with artificial intelligence, is that capture and review can now happen essentially at the same time. Instead of a human reviewing footage, AI-powered computers can analyze the images and make decisions about what's happening.

This new capability introduces new concerns for privacy. In this book, I've primarily been focused on still photography, but it's worth noting that AI technology is being used for video as well, and is evolving with many of the same concerns. In *The Dawn of Robot Surveillance*[24], Jay Stanley explores current capabilities and the privacy problems that can arise from both real and potential uses of the technology.

As the capabilities of cameras change and grow with the abilities of our computational systems and software, do the privacy expectations (need to) change as well?

How do we feel about a camera deciding when to snap the shutter vs. one being controlled interactively by a human?

Datastores represent another area for potential privacy concerns. If we use large sets of images to train various machine learning algorithms, what issues are introduced for the rights management of those images? Do we need clauses in our model releases to explicitly allow for a given image to be used for AI purposes?

Is This Good for the Company?

When we look at privacy concerns around our photos and how they might be used in the world of artificial intelligence, one area to be examined is the role of the enormous tech companies which collect vast quantities of data from their users, including photos. Companies such as Facebook, Apple, Amazon, and Google have massive stores of all forms of data, gathered in a variety of ways. But while these companies are similar in that they all are sitting on information originating with folks such as you and me, they take varying approaches to how they use, protect, or share that data. Here's a quick survey of the privacy landscape around these companies as of mid-2021.

Facebook

The best way to summarize Facebook and privacy is that they have a near-continuous history of abusing it.

Some of the issues have been widely publicized such as the Cambridge Analytica scandal[25]. Cambridge Analytica was a data research firm that created Facebook quizzes under the guise of academic research, but the quizzes collected a wide range of personal information from the users who participated. The quizzes also used a Facebook function to collect personal data from the *friends* of the person who took quizzes, which led them to accumulate a large amount of information for folks who had no idea that their information had been harvested. This data included names, birthdates, locations, and interests (such as page "Likes"). With the amount of information collected, Cambridge Analytica was able to develop detailed profiles about each person and then use those profiles for

advertising and other purposes. It’s known that the 2016 United States presidential campaigns of Ted Cruz and Donald Trump used the data directly, and it’s also been alleged that the information was used by Russian sources to interfere with the campaign. In Europe, the information is thought to have been used by at least two groups as part of the campaigns around Brexit.

Other Facebook privacy issues haven't always been seen by as wide of a mainstream audience, such as the story of Bobbi Duncan, who was inadvertently "outed" to her parents by the default sharing and privacy settings on Facebook, which caused them to disown her. As a result, she attempted suicide.[26]

Facebook cares only that you spend more time on Facebook, privacy be damned. They’ll pretend to care as long as it helps quell whatever current hot issue has them in the headlines in a negative light, but ultimately the company has shown repeatedly that your privacy is an afterthought. Unlike Google, which we'll discuss in a moment, Facebook continues to obfuscate the types of data they collect from their users, leading to widespread distrust.

One interesting photo-related privacy observation is that if you download an image directly from Facebook it appears that the metadata has been stripped… but if you perform an account backup, in addition to the images you’ll also get information files showing that Facebook has kept all the metadata (including location) behind the scenes in its databases.

There's a reason the Facebook Portal -- a device consisting of a home voice assistant coupled with a video camera and screen -- hasn't seen widespread adoption. Folks don't trust Facebook enough to want Facebook's cameras in their home, because folks have seen failure after failure when it comes to Facebook and privacy.

Apple

If we're talking about a spectrum of public to private with the mainstream tech companies, we have Facebook at one end, and Apple is likely at the other. There’s no other major tech company (especially one working in the artificial intelligence space) placing such a high value on privacy. Apple’s ability to focus on privacy is because their revenue model is different from the other big players in this space. Companies such as Facebook and Google have advertising as a key revenue generator, and the rates they can charge advertisers are dependent on the amount of personal data they can gather. The greater the amount of personal data collected, the higher the rates for advertisers. These higher rates are because with more data, it’s easier for advertisers to narrowly target their advertising to the individuals most likely to become their customers. With Apple, where advertising to consumers is *not* a core part of their business model, Apple doesn't want or need to have all of the details about your personal life. Instead, Apple focuses on revenue through hardware and services sales and designs its software such that it provides as much benefit to the end-user as possible without compromising privacy concerns.

What does this intense focus on privacy mean for everyday users of Apple products? It means that Apple sees privacy as a feature, and having privacy won’t hold the company’s users back from also having first-rate forward-looking technology.

Let’s have a short diversion into mathematics and statistics since it plays into how Apple discusses and implements user privacy.

Differential Privacy

We ought to take a look at a concept that’s front-and-center in Apple's privacy discussions and marketing. Apple makes it well-known that a core tenet of their privacy mechanisms is *differential privacy*.

Differential privacy is a mathematical definition of privacy where the level of privacy can be measured based on the interaction of multiple factors. A full explanation of differential privacy is beyond our scope (and can get *quite* technical), but it makes sense to at least take a basic look at how it works.

As Apple describes it:

> *Apple has adopted and further developed a technique known in the academic world as local differential privacy to do something really exciting: gain insight into what many Apple users are doing, while helping to preserve the privacy of individual users. It is a technique that enables Apple to learn about the user community without learning about individuals in the community. Differential privacy transforms the information shared with Apple before it ever leaves the user’s device such that Apple can never reproduce the true data.*

As Apple notes, differential privacy isn't their invention, but they've applied the concepts of differential privacy and used them for various features in their modern products. The company doesn't necessarily need to know all of the details of which data came from which user, but it wants to understand trends and patterns in the data. Some of the data that comes from customers is fairly benign. Some of it can be sensitive (such as the location of their home, or photos of their children).

Differential privacy allows one to apply mathematical values to quantify just how secure or anonymous the data is in a large data set. In this case, anonymity means that a human analyst (or a computer algorithm) working with the data is unable to identify individual persons. To make the data more anonymous, some noise is added to the data. The noise in many ways is a lot like the pixelization of a portrait. As we add more noise, the face becomes more anonymous. But as we add more noise, the portrait also becomes less useful because we can no longer identify who is in the portrait.

This measurement of "how much noise" leading to "how anonymous" is controlled using a parameter called epsilon (ε).

By using random noise in the data for anonymization, each time the data is queried, the level of anonymization decreases a bit. This happens because the aggregate results can be used to reconstruct the data. Averaging the results essentially filters out the noise. With larger values of epsilon, you have more privacy, but less accurate results from data analysis. Smaller values of epsilon give you more accurate query results at the expense of privacy.

Given that the amount of privacy decreases with each query, your epsilon value is directly related to how many queries are possible before the data set reaches a point where it can no longer be considered anonymous. This is referred to as the "privacy budget" of the data set.

Beginning with iOS 10 in 2016, this privacy "noise" was introduced into user data on the mobile devices *before* the data was sent to Apple. The data is then sent to Apple in a form that isn't directly linked to a specific individual. By aggregating these datasets, Apple can identify trends and patterns even with the addition of the privacy noise. Apple has chosen an epsilon value that it feels adequately protects the privacy of users while still retaining enough information that Apple's aggregate queries can provide useful information to be used for machine learning.

While researchers have pointed out that Apple's differential privacy isn't flawless, it's notable because it demonstrates a concerted effort to respect user privacy while also pushing forward with machine learning.

Local vs. Cloud

The use of differential privacy on its phones, tablets, and computers to better anonymize data before it is sent to Apple is one thing that sets the company apart from many others. In addition to the privacy implications of processing more data on-device than other companies, Apple's neural processing power in their phones and tablets means that their AI-assisted software has a performance advantage. Instead of having to send data to the cloud and then wait for it to return to the device for user interaction, the device can perform the needed calculations with built-in processing chips and provide near-instant feedback to the user.

It could be object recognition when sorting through your photo library. It could be a video editing program where you tell it to find the part of the movie with a car going into a tunnel. These sorts of operations will be faster without waiting for data to be sent to a remote data center, processed, and then returned.

Amazon

When we look at Amazon and privacy, there are enormous implications for the consumer data world, with Amazon being a major retailer of... everything. In addition to the physical products available for purchase, their media services and hardware also carry privacy baggage that needs to be handled. From a photography standpoint, we find ourselves most affected by a couple of different Amazon scenarios:

- Amazon's cloud hosting services, used by any number of major photo-related websites and companies such as SmugMug.
- Amazon's direct-to-consumer photography offerings with Amazon Photos.

In either case, the privacy story is a good one: unlike some of the other cloud photo hosts, Amazon doesn't mine user-uploaded photos for any purpose. That said, Amazon isn't without controversy when it comes to AI quality and images as we see when we look beyond their consumer services.

Rekognition

Amazon offers a service called Rekognition as part of its Amazon Web Services suite of software available for use by software developers. Rekognition is a service that analyzes images or video files for various information and is most often used for facial recognition and comparison. The controversy of the AI-powered Rekognition service generally falls into two broad areas of concern:

- The software is not without flaws (more details below), and
- Despite those flaws, it's being actively marketed and sold to law enforcement agencies who are making decisions based on its output.

We've looked at various examples throughout this book of photo-related AI as an emerging field. While there are some areas where the technology is solid and the results are reliable, there are also other areas that are still clearly in the proof-of-concept stage. A 90 percent success rate seems pretty good, but we can't tolerate 10 percentage points of error for many (most?) applications.

In 2018, several problems were discovered with Rekognition. Among them:

- The software disproportionately makes significant errors in the identification of women and people of color.
- Amazon has no oversight as to whether the software is being used following best practices for accurate results.
- In a test comparing portraits of members of the United States Congress to criminal mugshots, it misidentified 28 members of Congress as criminals. There's a joke to be made here about that number being low. [27]

As photographers, we hope that the benefits of AI-driven technology will assist us in making more captivating photographs and becoming more efficient with our work. As members of society, we ought to also look at how AI is being applied to photographs and insist that it be handled in a way that is just.

We want to apply the robot superpowers for good rather than evil.

While not explicitly acknowledging flaws with the software or how it was being used, in response to the George Floyd protests in 2020, Amazon announced a one-year moratorium on the Rekognition service being used by law enforcement.

Google

Google is known as one of the leaders in cloud computing because they've been successful at harnessing the power of large numbers of computers to fuel their AI and machine learning efforts. One of the reasons that Google's search engine continues to be so successful is that it can learn from the large volume of searches performed by users, and uses the user search behavior to feed back into algorithms powering its search rankings to better tailor them to how users search.

Whereas Apple's stance on privacy boils down to "your data is private and we don't ever want it," Google lands in a very different place, but is generally honest about it. Google collects a significant amount of data from users and processes that data in its cloud platform using various algorithms including artificial intelligence and machine learning. While Google promises to keep users' data relatively private, they do make it clear that they will use the data for their purposes. What data? In the context of this audience, let's talk about your photos.

Google Photos (as a product) offers cloud backup and sync of your images as a core feature. In addition to giving you an offsite backup of your photos, it also means the images are uploaded to Google's servers where they can be analyzed for various purposes. What sort of purposes? Many of the same things we've seen possible elsewhere:

- facial recognition
- object recognition
- image organization based on date and location metadata
- automatic image organization (albums) based on topic or subject

These features are useful, but unlike Apple where these sorts of algorithms run locally on your phone, tablet, or computer, Google's services run in the cloud.

What's the downside? You're giving this information to Google, who maintains it in a database along with everything else they know about you. Where you've been using Google Maps. Your web searches. Your travel

plans they culled from your Gmail account. One strength of Google's services is that they can tie all of this together for a very personalized experience. But that experience comes at the cost of one company having a *lot* of data about you. Google uses that data for advertising, with an extensive network of websites across all genres on the internet running code to tie together your behavior and target you for various purposes.

Google says its AI is better than competitors because it has more data than they do. And Google is probably right. Is the loss of privacy worth it for better AI?

Metaprivacy

Let's jump back to the topic of metadata and look at how it relates to privacy in an age where the metadata has world has shifted significantly due to artificial intelligence.

(If you skipped the previous chapter on metadata, now would be a good time to jump back and read it)

With traditional metadata, most of it was applied manually by a photographer, making it "private by default." There wouldn't be a bunch of telltale metadata unless someone chose to add it. The one piece of metadata that occasionally proved problematic from a privacy standpoint was geolocation data, usually because it was added automatically by a camera or smartphone GPS and not removed by the photographer when sharing the photo.

With artificial intelligence being used to deduce various pieces of metadata about an image, the metadata privacy situation changes. Instead of only having to worry about adding too much metadata to an image on purpose, we now will have concerns that AI algorithms will potentially pull information from a photo that was intended to be kept private.

All that information we discussed that could be inferred from a photo and could be used for various helpful purposes? We need to also remember that the same information could be used for harmful purposes.

With web search engines, conventions formed where a website owner could provide information about a particular web page (or set of web pages) and indicate that they shouldn't be part of a search engine's index. Essentially a site owner can put up a sign that says "Google, ignore this." Might we eventually see a similar bit of metadata on images to instruct AI to avoid scanning or interpreting the photograph? The request to not index web pages is a voluntary thing… there's nothing that *forces* a search engine to respect the request to not index a page. Presumably, the same sort of restriction would be in place if we request that an image not be scanned or used for AI learning.

Well-behaved algorithms would respect such an indicator… but algorithms are only as well-behaved as the individuals who design them, and not all individuals are well-behaved. Profit is a very strong motivator.

AI vs. AI

Let's take privacy and artificial intelligence to the next level. Let's use artificial intelligence in an attempt to maintain the privacy that can be otherwise lost through artificial intelligence.

Wait… how does that work?

If we know how AI learns from images in order to identify objects, people, and other things that one might want to keep private, one could potentially disrupt that learning. In 2018, researchers at the University of Toronto built just such a thing. They created software that can make small changes to an image with the purpose being to disrupt and derail the identification algorithms typically used to recognize faces and other objects in photographs. Think of it like an Instagram filter that makes changes that are barely perceptible to the human eye, but instead alters the data enough to prevent object and facial recognition technology from being able to correctly figure out what's in a photograph.

Now imagine you take large datasets, and start having these two algorithms work against one another. As the recognition algorithms get "smarter," the privacy algorithms adapt as well to make it harder for something to be identified. As I read about this study[28] it was described as an AI arms race, and that seems like an accurate representation. By pitting these two pieces of software against each other, they both become more adept at their

goals.

The folks from the University of Toronto study were looking at making their privacy-enhancing filter available for general use. I haven’t heard if that’s happened. I could see it either being offered as a standalone piece of software that someone could use if they were concerned about privacy and scraping of an image, but it also seems like a prime candidate to be acquired and incorporated as a feature in other services and software. That could mean a Photoshop or Lightroom plugin, an Instagram feature, or as an option when sharing images on your favorite social network.

23. https://www.techdirt.com/articles/20190707/18134542534/london-metropolitan-polices-facial-recognition-system-is-now-only-misidentifying-people-81-time.shtml

24. https://www.aclu.org/report/dawn-robot-surveillance

25. https://en.wikipedia.org/wiki/Facebook–Cambridge*Analytica*data*scandal*

26. https://medium.com/@monteiro/designs-lost-generation-ac7289549017

27. https://thenextweb.com/news/amazons-facial-recognition-ai-confuses-politicians-with-criminals/

28. https://www.sciencedaily.com/releases/2018/05/180531114620.htm

12. What is Photography?

With a look at what’s currently possible and as we look forward, we've seen that AI and related technologies are upending aspects of photography. What we previously thought to be true about cameras, photographs, and photographers is evolving in a world where we spend less time controlling the machines, and the machines spend more time controlling themselves.

What is a Camera?

As shown numerous times earlier in the book, the definition of a camera has changed. What was once an expensive heavy object that was only affordable and operable to someone with large amounts of money and skill is now accessible to nearly everyone in a first-world country.

A traditional SLR operated in a fully manual mode. And then it gained various "auto" features, making it easier for photographers to capture images as the camera could handle some basic functions of metering for exposure. The number of automatic features increased as we moved into digital point-and-shoot cameras and DSLRs. And then the smartphone came along and changed things again. First, it put a basic low-quality point-and-shoot camera into our pockets. That basic camera has evolved, and today's high-end smartphones contain image capture and processing technology that's on par (or even more advanced) than that of dedicated photography equipment.

John Gruber, the publisher of the Apple-focused website Daring Fireball, offered these thoughts after having worked with the camera in the iPhone 11 just before its release in September 2019:

> Several times over the past few years, I’ve had conversations along the lines of, “*I know they’re never going to do this, but wouldn’t it be cool if Apple made a real camera?*” As the iPhone camera system evolves, I’m starting to think Apple *is* making a real camera, right under our noses — or perhaps better said, right in our pockets.

John Gruber on Apple making a "real camera"

We will continue to see bigger cameras (DSLRs and mirrorless interchangeable-lens bodies) for the foreseeable future, but the traditional thinking where they were considered the “real” cameras and smartphones were second-rate options only useful in very limited situations is a mindset that is going away. Our definition of what we consider a camera has broadened and evolved to include new devices.

Who Is Still Buying Traditional Cameras?

With AI-infused smartphones in our pockets that can capture realistic nighttime images, automatically create sharp photos of moving objects, and other advanced features, where does this leave the market for more traditional cameras? What does it mean for DSLRs, digital point-and-shoots, or the mirrorless cameras whose popularity started rising around 2015?

In short, the traditional camera market is rapidly shrinking.

Anecdotally, most photographers wouldn't be surprised that point-and-shoot sales have dropped dramatically as smartphone cameras became "good enough" for the everyday person. But many folks would be surprised to know that DSLR and mirrorless camera sales are now on the decline as well. According to research from the Camera & Imaging Product Association, through 2020 there was a notable drop in *all* camera sales, to a level just around 50% of that seen in the previous year. With photography being more popular than ever among ever-growing swaths of the population, that camera sales decline certainly isn't due to a lack of interest in making pictures. And before we attribute this decline solely to the COVID-19 pandemic, note that sales dropped significantly in 2019 as well. It’s a clear trend.

The camera sales decline is because smartphones are now "good enough" for not just the most casual hobbyists, but for some fairly dedicated photographers as well. Even as a pretty serious photographer, I routinely go out for casual photo excursions or some travel and pack nothing in photo capture gear beyond my iPhone and a couple of accessories. Gone are the days of a backpack of equipment to make great vacation photos.

We once had a scenario where most of the serious camera purchases were for professional photographers. We're reaching that scenario again... not because non-professionals don't want or need good cameras, but because they already have them in their smartphones.

What is a Photography Company?

As we expand our minds around “what is a camera?” and “what is a camera company?” we also should consider an expanded mindset around what is considered to be a photography company.

Traditionally we might think of:

- camera manufacturers
- lens companies
- film companies
- paper and printing suppliers
- lighting equipment companies

We need to look beyond what we might think of traditionally. One could argue the biggest photography company right now is Apple, with the cameras and software on their iPhone being both crucial to their business and crucial to photographers of all varieties. If we look at what’s critical to photographers, we need to also consider software companies such as Adobe or Skylum. We ought to look at the hardware manufacturers who create the chips used by cameras for the capture and processing of our images. We need to consider that in an increasingly technical world of photography, various companies that might not have previously been on our radar should be considered as key players in our photography industry.

13. Photography Not for Photography's Sake

If you're reading this book and you're a photographer, the sort of images you make probably fall into a couple of broad categories. You might be a hobbyist who makes photos of their family, vacations, landscapes, or something else that catches your eye. You're doing it purely for fun. Or you could be a professional photographer who uses their skill with images to create photos for clients, whether they be portraits, commercial, photojournalistic, or another genre of work.

There is a different side of the photography world that we don't often consider. I'm referring to photography that isn't necessarily about making a beautiful image, but rather it's a utility to enable us to see what isn't possible or practical to see with human eyes. Photography is a tool, but the resulting images aren't the end work product. Let's explore just a bit of this world and consider the implications of artificial intelligence for this more utilitarian work.

AI for Accessibility

Those who have hearing impairments often depend on closed captioning for television programs, movies, and online video to help them understand what is being said in an audio environment. Automated closed captioning services are an efficient way to make information more accessible. Those with visual impairments can be challenged when viewing our photo-centric publications and websites, and artificial intelligence can help here as well.

Whether web pages are designed by hand or with a content management system such as WordPress, the result is HTML code specifying the site's text, layout, and images. For decades, HTML has supported the ability to include alternate text with each image. This text should describe the image, and as such, it can be used by the screen reader software used by visually impaired folks to understand what's in those photos that they can't see. Although HTML can include this alternative (or "alt") text, most images on the web don't have alt text specified, meaning that a significant portion of the internet is of limited use to those with visual disabilities.

Since the alt text should describe the image, this is an area where artificial intelligence can power object recognition algorithms to assist in making photos more accessible to all. I envision a couple of possibilities here, one of which is already happening:

- Microsoft has released an algorithm as part of its online Azure services and through a consumer app called Seeing AI that can describe what's seen in photographs. By using the Microsoft algorithm, one could automatically generate alt text for images on their website and publish the alt text as they publish the photos.
- Another possibility would be that the AI photo recognition capabilities could be built into screen reader software such that the image recognition is done "on the fly" even when alt text might not be present. While this relieves website publishers of the need to publish alt text for accessibility purposes, the publishers would miss out on some of the other benefits of using alt text (such as helping their search engine rankings). This solution also shifts the burden onto the site's viewer to have and use software with this capability built-in, but if one is already using accessibility technologies, perhaps this wouldn't be overwhelming.

Beyond the online world, there are at least a couple of groups pursuing the use of AI to assist visually-impaired folks throughout their days.

In 2020, Google developed a system to use an Android phone's camera to help guide a blind runner along a guideline on a racing course (such as a 5K)[29]. The phone tracks the movement along the course and warns the runner when they start to veer away from the guideline.

In early 2021 a group of researchers at the University of Georgia created a backpack that houses computing power attached to a 4K camera with AI processing power[30]. The camera is worn either on a vest or forward-facing fanny pack and communicates with the backpack computing device. A Bluetooth earpiece allows a visually impaired wearer to understand where they are and perform functions such as reading traffic signs or warning the person that there is a curb in front of them.

Regardless of where the application of AI occurs, this is a great use for image recognition and object identification to make the world more equitable.

Seeing What We Can’t See

Surveillance cameras aren’t new, and we’ve touched on a few concerns around applications of them earlier in the book. One interesting new application is for using AI to see things that aren’t visible to human eyes when making observations.

A study released in late 2019 by seven researchers at MIT[31] looked at the application of cameras and AI to see things and make conclusions based on the shadows cast by objects that are otherwise out of sight. Imagine a scenario where you’re looking down a hallway and there’s a junction to another corridor. The human eye can’t see down the side corridor because it’s around the corner and is blocked by the walls. However, whatever is down that corridor may cast shadows that might be visible. The study used those shadows to “see” what was around the corner.

This is a pretty far-off application of the technology at this point, but if we consider some of the utilitarian applications, such vision around corners could be useful for purposes such as self-driving cars, search and rescue, or military purposes.

Seeing Assembly Lines

Cameras are used extensively in various industrial and manufacturing situations to monitor the progress of automated tools. Software improvements including artificial intelligence are coming to these cameras as well, which will lead to increased manufacturing capacity and overall increases in quality. Much like with other areas of AI, research is being done in various areas and over time we’ll see cameras that can more accurately and reliably identify manufacturing issues than human quality checkers.

Much like we can use machine learning to train a computer how to identify objects in a photo, similar training data could be used to help a manufacturing or assembly line camera learn to verify correct operation, spot anomalies, or make decisions on what to do based on the state of the manufacturing process.

Beyond our photos for the photography industry, photo-based AI technologies will be applied elsewhere with interesting implications.

29. https://www.engadget.com/google-project-guideline-blind-vision-impaired-accessibility-202659152.html

30. https://www.engadget.com/ai-backpack-system-visually-impaired-intel-university-of-georgia-150025561.html

31. https://arxiv.org/abs/1912.02314

Ahead

Given where we've come from, and where we are now… what's next?

14. Varying Advances from Varying Companies

We've explored quite a few technologies that affect image capture, editing, and management.

If you look at the various companies we've discussed, different names come up in different contexts. Some (such as Apple and Google) are in many places... their work affects how we capture images on our smartphones, how they're processed and edited, and how we manage the collections of our work. Other companies are more specialized. Adobe is investing heavily in their Adobe Sensei AI technologies that bring new and improved capabilities to Photoshop and Lightroom both on the computer as well as mobile devices. Skylum’s Luminar has gone all-in on AI, even naming a recent release of their flagship software as “Luminar AI.” Other single-purpose apps are using AI for a specific function.

With hardware manufacturers, we can expect to see AI-powered improvements in cameras and lighting systems from companies such as Sony and Canon.

Different companies will bring different things to the table... the growth of artificial intelligence in the photo world isn't a single-company or single-platform effort. Together, the innovations coming from numerous hardware and software entities will mean an overall growth in our photography capabilities. We all benefit from the broad adoption and growth in AI-powered capabilities in our photography toolkits.

15. Responsibility

As we've seen throughout the book, artificial intelligence capabilities are becoming more and more powerful, with the ability to impact all aspects of photography. Capture, editing, management, and publishing our work is shifting further away from manual effort by humans into a world where much of that effort is now handled by powerful computing technology. But as noted by Stan Lee, with great power comes great responsibility. As a technology, AI doesn't take responsibility for itself. The responsibility is placed upon the humans that design, program, and use AI-powered systems.

I've often quipped that the best thing about computers is that they will do exactly what we tell them to do, and the worst thing about computers is that they will do exactly what we tell them to do. This applies to our AI programming as well. When technologists and programmers sit down to design and build the machine learning algorithms and artificial intelligence technology of today and tomorrow, they bring their past experiences. These experiences generally include both recognized and unconscious biases.

A computer can make decisions and "learn" to accelerate our prejudices.

How do we ensure that we use our AI powers for good purposes in responsible ways?

Moving Fast and Breaking Things

Let's talk about Facebook and its track record of responsibility (or lack thereof).

Facebook gained notoriety for its motto of "Move Fast and Break Things." This provocative slogan was originally viewed as a whimsical mantra… a nimble trait allowing it to experience rapid growth in both the number of users and features. Nobody ever claimed it was perfect, but folks generally chuckled about the motto as Facebook plowed forward with feature after feature. There were some missteps along the way, and Facebook occasionally had to backtrack a bit, but for many years it was growth on top of growth.

Around 2017, perception started to shift. A series of incidents rose to the level of making news in the mainstream media where Facebook got caught acting badly. Some of these were relatively straightforward situations where Facebook wasn't honoring privacy settings as advertised. Others were a bit more nuanced, including the fact that even when folks had chosen to make information private, Facebook was using that information for targeted advertising.

Breaking Things with Privacy

Another example of Facebook moving fast and breaking things was in decisions around what activity on the site would be private, and what would be shared with others. I referenced this story earlier, but as told by Mike Montiero in Design's Lost Generation:[32]

> *Bobbi Duncan was "accidentally" outed by Facebook when she was a college freshman. When Bobbi got to college she joined a queer organization with a Facebook group page. When the chorus director added her to the group, a notification that she'd joined The Queer Chorus at UT-Austin was added to her feed. Where her parents saw it. Bobbi had very meticulously made her way through Facebook's byzantine privacy settings to make sure nothing about her sexuality was visible to her parents. But unbeknownst to her (and the vast majority of their users), Facebook, which moves fast, had made a decision that group privacy settings should override personal privacy settings. Bobbi was disowned by her parents and later attempted suicide. They broke things.*

Some mistakes are easily repaired. Some, like the example noted by Montiero, have much more dire consequences that aren't reversible.

I point out this example not because it's a use of artificial intelligence, but rather because it's an example of where decisions made by humans impact how computers and privacy will interact. Humans from those same organizations then guide and direct its use of artificial intelligence.

Breaking Things with AI

Facebook made it clear that they were engaged with artificial intelligence experimentation, and they ran into a significant problem in late 2017. In their AI research, they'd developed chatbots to communicate with each other and "learn" based on responses. At some point, engineers noticed that two of the chatbots had developed their own language and were communicating with messages unable to be understood by the researchers. The bots were originally developed to interact and negotiate, and they morphed into a sophisticated code where their negotiations got fierce, including the bot "pretending" to like something to sacrifice it later in the negotiation.

At the point where the research engineers realized the bots were communicating in a meaningful code, but one where humans couldn't understand the intention or consequences, the experiment was shut down.

If we look at Facebook's poor record of decision-making when they embrace their motto of moving fast and breaking things, we could conclude that they're moving too fast without considering the repercussions of their decisions. We've talked about how artificial intelligence often allows us to move faster. Can we trust a company like Facebook to move even faster and break more things, given their history of moving at human speed and breaking plenty of things?

Like Human, Like Machine?

Let's step back from Facebook for a moment. One of the goals of artificial intelligence is generally to increase our technology power to the point where the computer can act and make decisions on par with a human. And when we're talking about things like object recognition or cloning an item out of a photograph, that sounds great. Bring it on. Let's make our lives easier.

But given our human track record of making bad decisions, some of which have significant negative consequences on the lives of other people, there's a question to be asked if "as good as a human" is good enough when it comes to some uses for artificial intelligence. If Facebook's humans have moved fast and broken too many things, can we trust those same humans to design and program AI solutions to do the same?

Will we hold artificial intelligence to the "as good as a human" standard, or will it be held to a higher standard? If we want AI to make the world a better place, it seems a higher standard will be needed.

We're seeing discussion around this topic when it comes to self-driving cars. We have a long and well-documented history of the number of mistakes made by humans that lead to accidents with injury or death. Is it acceptable for self-driving vehicles to have a similar number of errors, or do we hold the computers (and the folks who design them) to a higher standard?

As Biased as Humans?

Another factor we can't ignore is that our first iterations with various AI technologies have often produced results perpetuating various (sometimes subconscious) biases.

A late 2019 study by the National Institute of Standards and Technology (NIST) found that facial recognition systems misidentified persons of color up to 100 times more often than white individuals. This study looked at a variety of systems with varying levels of discrepancies discovered in the recognition of white versus non-white faces. It's unlikely that any of these differences are the result of an intentional decision on the part of anyone involved with the development of the various recognition algorithms, but unconscious and unintentional bias can find its way into the machine learning processes. Recall from our discussion of how most machine learning works by "learning" from various training data sets which are used to help it classify information and make decisions. The type (and quantity) of information used for training data will have a significant impact on the results of machine learning. Basic mathematics can show how minorities will be underserved by artificial intelligence unless care is taken to ensure they are well-represented in training data scenarios.

Let's consider facial recognition accuracy.

Imagine that you're training an algorithm with facial recognition capabilities and you have a sample data set of 10,000 images that you're going to use to train the computer to make decisions about how to classify the folks who are recognized. For simplicity's sake, let's imagine the data set only contains white, black, and Asian individuals.

And let’s say that 8,000 of those images are of white individuals, with the other two being split evenly between black and Asian people.

Ignoring all other factors, if you use this as a training data set for facial recognition, you’ll have trained your algorithm with eight times as much data about white folks as data about either of the other two races in question.

One would expect that the resulting AI algorithms will be much more accurate when classifying white individuals, if for no other reason than the fact that there was a vast discrepancy in the amount of training data for the different races. Even if the split between the races was representative of a given population, the significant differences in the amount of data used for training are problematic. Without additional training data provided for the minority races in this data set, the algorithm will make many more mistakes in identifying non-white individuals.

Let’s tie this back to photography. Facial recognition is applied in our cameras (for eye and focus tracking) as well as in our image management and editing programs. We can use AI software features to help automatically identify folks in our images. Algorithms that enhance and correct images use their “knowledge” of facial features to make portraits look as good as possible. If the software is more accurate and more effective when working with folks of one race versus another, we end up in a scenario where we introduce disparities in the quality of our work and aren’t able to serve our clients equally.

Computers, for better or worse, do exactly what we tell them. If we want them to be effective when recognizing and working with folks of diverse backgrounds, we need to ensure that the software engineers building our algorithms use accurate and equal data sets for all populations and take active measures to eliminate unintentional biases.

When Private AI Meets Government Use

There’s at least one company that’s carefully walking along the line of what’s legally and socially acceptable when it comes to using artificial intelligence for questionable purposes. Clearview AI is a facial recognition software company providing services specifically for law enforcement purposes, enabling various police agencies to match names to photographs of faces. Founded in 2017, Clearview has scraped the internet, collecting over three billion facial photos. The company operated without much public notice until a January 2020 New York Times article titled “The Secretive Company That Might End Privacy as We Know It.”

Clearview’s images have been scraped from social networks such as Facebook, Twitter, and LinkedIn. The source images are available to the public, although scraping for these sorts of purposes typically violates the terms of service of the social networks. At least three companies have sent cease-and-desist letters demanding that Clearview stop the practice. As of early 2021, it remains unsettled whether Clearview’s methods are illegal.

Aside from questions about the sourcing of their data for the large facial recognition database, other questions exist around how the company markets its services and whether the use of such a database by law enforcement is a responsible use of the technology. Given what we’ve already discussed around the problems of bias in machine learning when used for facial recognition, putting those tools in the hands of government agencies bears scrutiny. If a social network mis-tags one of your friends in a photograph, it’s not a big deal. But if police departments (with the ability to arrest folks and cause long-term impacts on peoples’ lives) are given unlimited access to facial recognition databases, there can be a lot of unintended consequences.

Consider what might happen if a crime is committed, and there’s a photo of the suspect to be used to try to apprehend the criminal. Maybe the suspect’s face was visible on a store’s surveillance camera. Or maybe someone saw the crime being committed and captured it with their smartphone. We spent a lot of time talking about how great our smartphone cameras can be, and if something goes sideways, bystanders are likely to take note. Regardless of how it was obtained, the police now have an image of the suspect.

If the police have access to a facial recognition and matching database, it should be trivial to feed that image into the software and let the algorithm identify who is their suspect… right? But of course that all depends on the accuracy of the algorithm. And if we’re talking about the police naming a suspect and making an arrest, the algorithm needs to be correct. Close doesn’t count. It’s not good enough if the algorithm is “usually” right or only “sometimes” misidentifies someone.

If we’re going to be arresting people based on AI, the software needs to be right nearly every time. 90% accurate doesn’t cut it, and as mentioned earlier, the actual accuracy rate for face-detection cameras in use currently is much

lower.

Compounding the fact that we have less-than-perfect technology potentially being used to arrest individuals is the fact that Clearview is incredibly secretive about its customers. Law enforcement agencies are reluctant to admit using the system, and there have been multiple instances where it's been discovered that an agency was using Clearview, only to stop doing so once it became public.

If the software truly were accurate, and it was only being used for lawful and reasonable purposes, why have we not seen either Clearview or their customers willing to discuss the use of the system? The reality is that it's quite problematic. Facial recognition software with a significant error rate in the hands of police organizations that already have issues with racial equity is a recipe for a lot of innocent folks being arrested due to faulty systems and processes.

In mid-2020, after Canadian authorities announced a privacy investigation into Clearview AI, the company decided to exit the Canadian market. As they say, where there's smoke…

Meanwhile in the Private Sector…

While governmental use of questionable AI technology is quite concerning (since the government is granted numerous special legal powers over the lives of its residents) there are also examples of private companies employing artificial intelligence in dubious ways.

Over eight years, drugstore chain Rite Aid deployed facial recognition technologies at many of its stores, primarily in the New York and Los Angeles areas. In mid-2020, Reuters investigated the situation and discovered a variety of concerns.[33] The system used software from a company with ties to the Chinese government. When looking at the stores where the cameras were deployed, it found that the cameras were in use much more frequently in poorer neighborhoods and neighborhoods where the population is less likely to be white.

Rite Aid's stated goal of the program is to deter theft. In theory, the system could be used to provide an alert when the visible cameras detected someone in the store who had previously been identified as a shoplifter or otherwise not allowed to be on-premises. Once the system identified an individual, it could alert loss prevention staff to make contact with the customer. As we've discussed previously, the challenge is that facial recognition isn't foolproof, and existing technology is notably less effective in matching the facial features of non-white individuals.

While Rite Aid defended the use of the technology in comments to Reuters, the retailer also discontinued the entire program shortly before Reuters released the information publicly.

Rite Aid isn't alone. Other well-known retailers such as Walmart and Home Depot have also deployed facial recognition technology. And while these are big household names, smaller businesses haven't been exempt from AI-related missteps. In mid-2021, a 14-year-old black teenager was denied entry to a roller skating rink near Detroit. A facial recognition system flagged her as someone previously involved in a fight at the venue. That system failed; the teen had never been to the roller rink before the day she was denied entry.

As we learn more about the technological capabilities of these systems, as a society we need to come to decisions around the impacts when unproven AI is used in ways that can affect the lives of those it targets.

The Software Education Ethics Problem

We've ended up at a place where software engineers often build software first and ask questions later about possible misuse. It's not hard to understand how we've gotten to this point if one takes a look at the path that many folks take in becoming a software developer.

Some software developers attended formal classes to receive training in their fields of study. They may have majored in computer science, electrical engineering, information technology, or another tech-related discipline. Although the course requirements vary from school to school, finding even a single class about ethics is a rarity in one's technical class load. As late as 2019, mainstream publications such as Wired ran articles noting that major colleges were working to begin adding ethics classes to the prescribed curriculum for folks majoring in computer

science and other technical fields. If these classes have just been added in the last couple of years, it follows that nearly all of the formally-trained software professionals at companies such as Google, Amazon, Apple, and Microsoft likely had little to no formal ethics training. They learned about computer hardware, and compilers to translate their human-readable code into machine language, but they never learned about considering all of the implications and effects their code might have on the world.

While many in the software industry have degrees in a technical field, there are also a significant number of developers, designers, and people in other software-related job roles who are self-taught or otherwise moved into their technical positions without extensive formal training. It's not uncommon for someone to begin in an entry-level technical position such as performing basic computer troubleshooting or handling inquiries in a call center before learning on the job and moving into higher-skilled roles. A motivated individual might teach themselves basic programming skills and work their way into more advanced roles. When someone is learning the skills they need to accomplish a given technical task, considerations around ethics and responsibility are rarely a concern.

Whether someone was formally trained (in a curriculum lacking in ethics topics) or learned independently, the result is that most of the folks working in technical roles around software development (including artificial intelligence) have not received ethics training. Considerations around privacy, malicious bots, censorship, or systemic racism aren't part of their history, and thus don't receive the amount of consideration that is arguably required. Computers are only as ethical as the developers who program them. As we move into a future where we trust automated systems to make more (and more important) decisions, a higher level of importance needs to be placed on various ethical considerations around how those systems are developed, tested, and refined.

An Increasing Consciousness

All hope is not lost. Although systems of artificial intelligence can be used for questionable purposes, the industry is currently playing catch-up to where it should be when it comes to ethical awareness and ensuring that future uses of technology are for appropriate purposes.

Companies such as Google and Amazon now routinely make improvements to their attempts to do better and to apply ethical standards to their use of artificial intelligence and similar technologies.

32. https://medium.com/@monteiro/designs-lost-generation-ac7289549017

33. https://www.reuters.com/article/us-usa-riteaid-software-specialreport/special-report-rite-aid-deployed-facial-recognition-systems-in-hundreds-of-u-s-stores-idUSKCN24T1HL

16. It's All Around You

We've looked at quite a few different applications of artificial intelligence for photographers throughout this book. We looked at how AI is being applied in cameras to help photographers capture the best possible image when the shutter button is pressed. We looked at how AI is in play even *before* the button is pressed in how it can assist with autofocus, face detection, and eye-tracking. After capture, AI can help us find images and otherwise manage our ever-growing collections of photos. Some estimates suggest there will be 1.4 trillion photos made in the year 2021.[34] AI will ease the pain of sorting through our images. When it comes to editing, advancements with AI are making it easier to improve our average photos and to add that extra bit of oomph to our best ones.

In 2021, we're at a point where a camera, smartphone, or piece of photo software having artificial intelligence is a notable feature. It's something that will get called out on the marketing website and you'll hear about it when the company demos its products on stage or at a trade show. As AI becomes more common across our photographic world, it will become less notable. Remember when something being "digital" was the hot new thing? In the future, AI will be baked into the products and won't be called out as broadly as it is today.

I've mentioned over two dozen companies so far with their fingers in various bits of the AI photography pie. For some companies, including numerous companies I haven't mentioned, their AI *is* their product and it becomes a component of the overall photography scene. For other companies, AI is just one feature off to the side that is used to help us make better pictures.

The widespread proliferation of photographic AI technologies is an industry-wide situation. Different companies are making advances in different areas as they relate to that company's products and services. Camera manufacturers such as Canon and Sony will focus on technologies for their camera bodies. Lighting companies such as Profoto will look at how AI impacts their strobes and speedlights. Image management software from companies such as Adobe or Apple will leverage AI to help us sift through and make sense of our vast photo collections. Vendors such as Topaz Labs, Skylum, and Phase One will continue to push forward with new AI advancements that help us edit our work to make it as impactful as possible and to allow us to perform our editing as efficiently as possible.

We won't end up with a few companies being known as the ones that "do" artificial intelligence, but rather we are likely to find it rare when a significant piece of photography electronics or software *doesn't* include the use of AI.

34. https://focus.mylio.com/tech-today/how-many-photos-will-be-taken-in-2021

17. AI-Generated Artwork

We've explored the state of how artificial intelligence is helping photographers during our photoshoots as well as when we edit and manage our images. But if we look at the capabilities of artificial intelligence and software that can learn about artwork, we also should consider the concept of photographic artwork being created on demand by the computer, with no photographer (or camera) involved at all.

The Face Place

Although there are numerous applications for using AI to *enhance* photos we capture of people in various situations, can that same machine learning be applied to *generate* an image of a person from scratch? Consider the amount of machine learning being done against a variety of images of human faces. A machine can learn how the various parts of a face come together to create the whole. Lips. Noses. Eyes. Ears. Cheeks. Facial hair including eyebrows, mustaches, and beards. A computer that can detect an eye blink and replace it with the open-eyed version of an individual could presumably create a new image of someone as well, couldn't it?

It can.

A couple of years ago, Nvidia (best known as a graphics hardware manufacturer) released open-source software that can be used to generate realistic human faces. Consider these two images of people that don't exist, but were generated by a computer to represent two hypothetical celebrities:

AI-generated celebrities. Photos courtesy of Nvidia.

Nothing obviously looks wrong to our eyes. Looks realistic to me!

Want to see this image-generation software in action? Point your web browser to ThisPersonDoesNotExist.com and a brand new person will be generated for you on the fly.

From Face to Scene

If we can generate faces, can we expand on that learning to generate entire scenes? Consider a scenario where you're sitting at your computer and need an image for a project. Instead of finding a location, driving there, perhaps bringing a model, ensuring the lighting is correct, and capturing an image with your camera, what if you could simply tell your computer to take care of the whole thing for you?

> *OK Google, generate a photograph of a teenage girl leaning on the railing of a boardwalk with the ocean in the background. She should have blonde hair, blue glasses, and be holding a red balloon.*

> *Hey Siri, make a photograph of a sunset scene with a silhouette of a tree without leaves on the right side of the picture and two people holding hands on the left side of the picture.*

The first iterations of such a venture will likely be rough. But like everything else in the world of computational photography, as processors get faster and algorithms get better, we'll see improvements over time.

We're not yet there for generating entire scenes, but work has been done in this area, and the initial efforts are proving that it's possible. In a joint research project between MIT and IBM, an online app called GANPaint Studio allows one to "paint" new objects onto existing images.[35] One example involves an image of a kitchen with cabinets along the wall and the software was used to "paint" a window in the wall among the cabinets. Basic object addition is a great first step to eventually being able to conjure entire computer-generated scenes.

Another effort in image generation comes from the Allen Institute for Artificial Intelligence. They've created a machine learning algorithm capable of generating basic scenes using only text input as a guide. As of late 2020, anyone can try it on their website. The service is quite rough.[36]

Here's an example of an image it generated of "three people on a blue couch":

Three people on a blue couch as generated by AI

And here is what we get when we ask for "two cars on a road":

Two cars on a road as generated by AI

It’s a start, but it’s far from fully baked.

If it becomes trivial to use computers to generate images of our own design, what does that mean for photographers, images, and human perception of the results?

RIP Stock as We Know It

Stock photography was once a viable niche in which a photographer could make a respectable income. Over the past fifteen years, the proliferation in prosumer cameras led to a correlated proliferation in folks looking to make a bit of extra money with their images, which led to the prices for stock photography plummeting. What was once primarily an industry of rights-managed stock images (where someone would license an image for a specific purpose and timeframe) has given way to a world of royalty-free stock photos (where the buyer can use the image for any purpose). As supply increased, prices fell, and at this point, almost all money in stock photography is coming from the microstock industry where photographers are making pennies per image sold.

It’s bad now for photographers wanting to make money with stock photography, but it’s going to get even worse soon.

The nature of stock photography is that the images are often fairly generic. It consists of images of generic people in generic scenes performing generic tasks. It consists of images reflecting a given mood, feeling, or idea. If we look back at the sort of things that can be recognized by artificial intelligence and our discussions earlier in this

book, and we then consider what could be created as AI-generated “photographs”, we find that the new technologies could probably replace many, if not all, categories of stock photography.

Want an image of a businesswoman on a telephone? Have the computer create one.

Need a couple of photos portraying a dismal, gloomy mood? Conjure some up via AI.

Want a photograph of a race car going around a corner of a speedway with some smoke coming from its tires? I suspect that your computer of the future will be able to create such a thing… and so will the computers of most clients who would have previously purchased stock images.

If our computers can generate generic images on demand, there will be no reason for companies to purchase that style of an image from a stock agency or website. And without anyone wanting to buy said images, the remaining income opportunities for stock photographers are going to dry up. For those of you still making any residual income via stock photography sales, enjoy it while it lasts.

What is Real, Anyway?

It’s often said that communication isn’t just in the message being sent, but also in how it’s received. I’ve always believed that photography is not just in the making of the picture, but also in the viewer’s observation and reaction to that image. If a photographer makes a picture of a tree falling in the woods, but never shares it, did it even happen? If we start altering the process of the creation of images and remove the human element, how will those changes affect how images are perceived by our human viewers?

We are already seeing the possibilities of the technology, and starting to grapple with the implications for what we believe as far as photography and reality.

The notion of an image being “photoshopped” to alter reality isn’t new… in fact, photographers have been accused of manipulating images in this way for nearly as long as Photoshop has existed. Folks have started to learn that they can’t believe every photo they see… but what about the next level?

Not Just Fake, but Deepfake

> *…your scientists were so preoccupied with whether or not they could, they didn't stop to think if they should. — Jurassic Park, 1993*

We’ve looked at dozens of possible uses for AI in the realm of photography, from capture to processing to management. One theme has been that in general, these changes help us move forward. They help us make a better photograph in our camera, help us prepare the final image, and help us manage our collection of photographs. In some cases, we are much more efficient than we were with older processes for similar functionality. In other cases, we have entirely new possibilities.

But what if one were to use the power of artificial intelligence, mixed with imagery… for less noble pursuits?

The term *deepfake* was coined in 2017 as a portmanteau of “deep learning” and “fake” to describe the process of synthesizing (blending) human images to create realistic output, generally as a video. Using machine learning, the result is that one person's image could be "transplanted" onto another, creating a resulting fake image to deceive the viewer. There have been a couple of themes so far in deepfake videos that have spread online:

- Some have been pornographic, where a well-known celebrity’s face was placed into an explicit video, making it appear that the celebrity was engaged in sexual acts.
- Others have been political, creating a video showing a particular political figure engaging in nefarious behavior or speech at the whim of the deepfake’s creator

With the ability to create deepfake videos becoming easier, the public skepticism around whether a still image has been edited with Photoshop will now need to expand to video as well. Deepfake videos can be edited or generated to deceive the viewer in the same way as can happen with a still image.

Spot the Fake…

The adage says that seeing is believing, but we're now in a world where that can't be our only test for proof. We're able to see plenty of things that are altered from reality, or entirely made up from scratch. The same skepticism that we have applied to potentially edited still photos is now warranted for videos as well.

Our ability as humans to detect when an image (or video) has been edited or when it should be considered "fake" is questionable at best. It's not always clear whether we're looking at reality, an interpretation of reality, or an entirely generated scene. Events of the past few years have also shown that intelligent folks won't necessarily trust another's opinion when they claim "fake news."[37]

If humans are flawed in our ability to detect when a photo or video has been manipulated, can we turn to computers for help? Technology got us into this messy situation... can technology help us get out of it?

It can.

Researchers are looking into various uses of technology to be able to spot edits in photographs.

One detailed paper released in mid-2019[38] specifically looks at the use of the warp tool in Photoshop. The warp tool allows a Photoshop user to "push" part of an image and is often used to make subtle adjustments to an individual's figure. Perhaps we want to alter the cheekbones a bit, or slim someone's hips. The warp tool allows for these changes, making alterations to those areas of the image and altering the area around them to maintain a realistic-looking photo. Unless one knows what the "before" version of the photo looked like, it can be hard to know that we're looking at a version "after" the edits.

For the study involving the warp tool, the researchers used large data sets (over a million images) specifically chosen to contain faces. Some had been edited via automated means, some had been edited by a professional artist using Photoshop, and others were left as-is. The study looked both at a binary yes-or-no "was this face manipulated" benchmark, along with attempting to identify *which part(s)* of the image had been manipulated. The study showed that the automation could clearly outperform humans in judging simply whether image manipulation existed, and quite frequently could identify the specific area of a photo that had been altered.

As artificial intelligence progresses to allow for the creation of deepfake videos and other versions of altered reality that might be used for malicious purposes, AI can also be used to counteract these efforts and to identify what is authentic vs. what is fake. I would expect to see this sort of technology be integrated into media platforms of the future. Perhaps when you watch a video on a social platform such as Facebook or YouTube, there could be a trustworthiness indicator providing an evaluation of how likely the material portrays reality.

In September 2020 Microsoft announced new "Video Authenticator" software with capabilities to analyze a video and provide a confidence score as to how likely the software believes the video to be authentic vs. manipulated. By analyzing various detailed areas of the image, the software notes whether it's likely that they were manipulated or whether it's an original capture. Because this software's output could be used by malicious folks to alter the methods of creating deepfake videos, Microsoft is only releasing the software to known parties as part of a broader initiative on responsibility in artificial intelligence.

When both good and evil folks have access to technology and can use it against each other, we again run into scenarios where the technology is neither good nor evil, but its uses are up to those who deploy it.

Can Fake be Good? It Can.

Not all applications of this technology are necessarily evil. Let's turn our thoughts to the world of film production. Some folks from Disney Research Studios released a technical paper that boils down to being able to take the image of any person's face and being able to animate it onto another person's body at over a megapixel of resolution. While the ability to create higher-resolution deepfake videos certainly exacerbates concerns about fake and deceptive propaganda, there can also be valid applications for entertainment purposes.

Consider the ability of a movie studio to create a character appearing to be a well-loved actor or actress from the past. What if they bring back someone who appears to be a classic character that's no longer possible? Consider seeing Carrie Fisher as General Leia Organa, or Alan Rickman as Snape. Moviegoers have an emotional connection to these characters, and while other actors *could* fill in as replacement characters, the ability to digitally bring back departed actors into the roles we loved (or despised) is compelling.

Fake is the new reality, but only if we let it become so.

35. https://ganpaint.io

36. https://vision-explorer.allenai.org/text*to*image*generation*

37. Whether one chooses to share or ignore evidence that something is false is an entirely different question altogether - it's not one of technology, but of morals.

38. https://arxiv.org/pdf/1906.05856.pdf

18. (Not) Automated Out of a Job

With the talk of AI, machine learning, automation, and other technology advances impacting the photography world, many photographers fear that computers are going to automate them out of a job. If we look back at the last thirty years of the photography industry, we've seen significant disruption related to technology. In the introduction to this book, I noted that we've seen two big shifts:

- The internet led to a shift in photo distribution from being primarily based on photographic prints to being based on digital files.
- Accessible, affordable digital cameras led to a shift which made ownership of high-quality photo gear within the reach of most individuals.

Have these technological changes led to many photographers seeing declines in various areas of their business? In many ways, they have. On the other hand, they've also opened up new possibilities for other revenue streams. Let's look at each one, and then also look at what AI might mean for the photography business.

Internet and Shifting Photo Distribution

As consumers of photography have gotten used to sharing photos on the internet (first via email, and then via text messaging and social networks), the market for casual photo prints has dried up. Photographers who used to make significant revenue by selling small prints have been hit hardest. Consider a photographer specializing in senior portraits… in the past, this photographer could've counted on most clients purchasing numerous wallets, 4x5, or 5x7 prints to be given to friends and family in town and mailed to the senior's relatives around the country or around the world. With the now-ever-present internet, these gift prints are much less common, and those folks who previously would've received a small print will often just browse those photos online.

On the other hand, photographers who have adapted to this internet delivery have been able to find profitable slices. With a consumer base that values instant access to online resources, a photographer who can deliver that as part of their service offering has an opportunity to discover new revenue streams. Photographers in the event, commercial, and sports photography genres have new opportunities to provide same-day photo delivery for their clients. Businesses that value these services are willing to pay for them. In my work as an event photographer, for a client who wants real-time image delivery, I have revenue opportunities that wouldn't have existed in a world where I was delivering prints a few weeks later:

- I can hire an assistant who works with me to edit and share images as I shoot. This is an income opportunity for the assistant and me.
- If my client doesn't have staff available (or would prefer that I do it), I can post images directly to my client's social media accounts. This is a billable service.
- For events where clients want their attendees to have access to the images, I can host an online gallery. This is another billable service.
- If those attendees want to purchase prints, I can sell them and make a profit on the sale.

Some older revenue streams dry up, and some new ones present themselves.

Digital Cameras Shifting Business

Before the development of affordable, easy-to-use digital cameras, accessibility of high-quality photography gear was often a barrier to the everyday consumer which led to revenue for photographers. There was a point where you needed to have expensive, complicated photography gear to make great photographs. Photographers could capitalize by having that gear and investing the time to learn how to use it. They solved a problem for a client by providing value in photography services that the client couldn't access themselves.

Then digital cameras came along, got affordable, and got easier to use.

Having the photo gear was no longer a selling point for many photographers. There are great cameras available today for a few hundred dollars at any big-box retailer that are capable of making photos as good as the ones made

by fancy professional-exclusive cameras twenty or thirty years ago. Consumers who have been willing to spend a bit of money on their photography hobby and spend time learning how to use their camera have been able to make some great images that were previously out of reach of their capabilities.

Photographers who provided services where they were only the keeper of the camera (without adding significant additional value for their clients) have found it hard to sell their services.

With consumers buying better cameras, we've seen an interesting trend in photography education. If we go back fifteen years, large numbers of photographers observed this trend and got into the education business. Those consumers who were willing to spend $500 or $1500 on camera gear were often willing to spend money on education to learn how to use that gear to make interesting photos. Successful photographers often became (at least part-time) educators and could make a decent income either teaching basic photography classes or leading photographic tours and workshops.

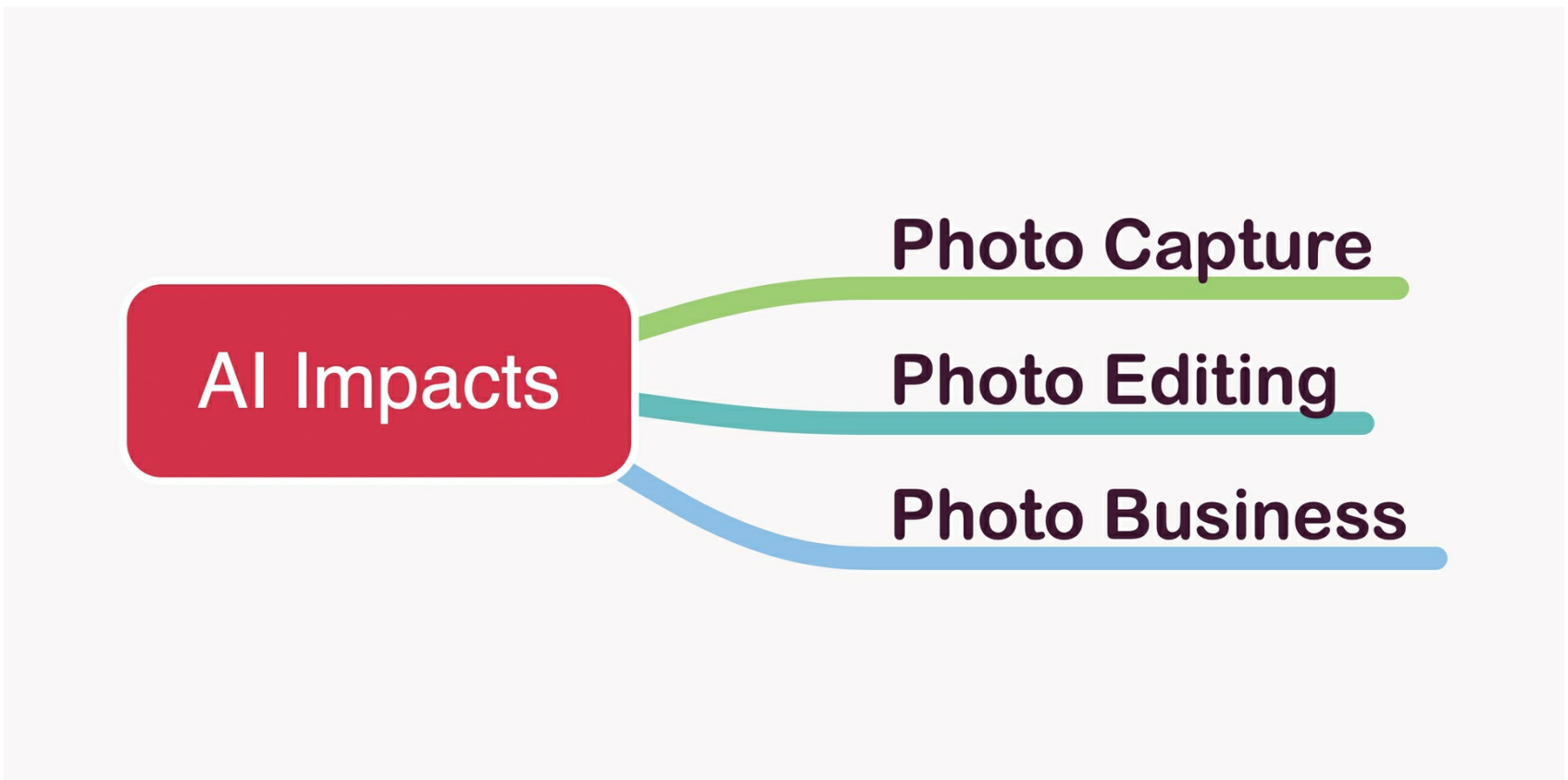
Like any market, the photo education market for consumers has adjusted and corrected over time. As smartphone cameras get better, we see folks choosing *not* to purchase a more serious camera, and they often rely on the camera they always have in their pocket or use instead of a more serious piece of photography gear. First-time DSLR buyers were often willing to spend some money on education to learn how to become proficient with their new tool. Those who forego purchasing a dedicated camera and instead choose to use their smartphone are less likely to spend money on education.

While the internet shift is likely a long-term one, the shift with affordable digital cameras is already changing as they're replaced by ever-more-capable smartphones.

The AI Shift in Business

I've called out the realm of artificial intelligence as being a third major shift in the photography world. Since it's the focus of this book, let's spend a bit more time considering what that means for the future of the professional photography profession. Will you be replaced by a robot?

The End-to-End AI Impact



Artificial Intelligence will impact all aspects of a photography venture

As we've explored throughout the book, artificial intelligence is impacting photographers in the three major areas of their ventures: image capture, image editing, and the business of images.

Are We Human?

Photography is hardly the first or only industry that is facing a changing world involving increasing amounts of computerization and automation. At times, this automation eliminates work previously done by human workers.

- Grocery stores have had self-checkout kiosks for several years, allowing shoppers to scan and bag their purchases. Whereas previously a cashier would serve one individual at a time, now a single cashier can oversee a half dozen kiosks. The cashier's role is no longer to scan and bag groceries, but rather to assist self-checkout customers when there are exceptional circumstances or a problem with the machines.
- Rapid-delivery e-commerce (such as that provided by Amazon.com) is enabled in part by augmenting the warehouse workforce with robotic units that can efficiently move products and packages to their destinations in the building. Instead of a person carrying a box or basket of products a hundred yards to place it onto a conveyor for packaging, a robotic unit can make that same movement. As products arrive at the warehouse or their location changes within the building, humans would need to look up the location information. Robotic units communicating wirelessly with a centralized product database will know where every product is at any given moment.
- We have automated transit systems in various locations. Whether it's a short-distance "people mover" to shuttle folks between terminals at an airport, or a broader citywide subway, the traditional role of a transit vehicle operator is no longer required when a system has a closed route and can be programmed for various scenarios.

What does this mean for the photography workforce, and what about the human element that might be lost when various jobs are automated? Automation of a role or task might lessen the time needed for a task, or create repeatability, or get us closer to technical perfection, but we should also consider that those aren't always our ultimate goals.

Is Perfection Perfect?

Over the past few years, soccer (football for non-US readers) introduced the use of a Video Assistant Referee (VAR) to assist the on-field crew with ensuring that various foul decisions were correct. The VAR role isn't fully computerized (a human referee is operating various cameras to identify questionable decisions), but it is an introduction of technology into an aspect of the sport that was previously entirely human. Until VAR, the on-field referees' decisions stood alone, and as with any sort of human decisions, some decisions seemed too close to call, with the occasional clear error. While a single referee decision in soccer doesn't usually have a direct impact on the outcome of a match, it does add a bit of variability and humanity into the game, much like the various actions by the players. Reaction to VAR has been mixed... while it has helped correct some clear errors, there's debate around whether the interruptions to the game are worth it in the cases where the judgment truly was a close call.

We come to the concept of appreciating imperfections. There’s a name for this… wabi-sabi, which is a Buddhist and Japanese philosophy that roughly encapsulates the notions of beauty in imperfection. It’s become a bit of a design trend, but we can also examine the underlying philosophy and consider how it may play to human emotions and desires in a way that won’t be eclipsed by artificial intelligence, at least not anytime soon. The concepts of wabi-sabi include an appreciation for the simple and unpretentious, as well as appreciating how things age over time and a recognition that the aged appearance often creates an interesting aesthetic.

Things created by humans, even when designed to be the same, vary a bit in their creation. Consider a portrait session for multiple subjects. Commercial photographers often are asked to create a series of headshots or staff photos for a company in a similar look and style so they appear consistent on the organization’s website or other marketing materials. Even when several folks are photographed on the same day, in the same studio, with the same lighting in use, there will be variations between the images. A head will be at a *slightly* different angle. There will be variations in the smiles. Natural differences in the faces lead to naturally different portraits, even with a similar pose and lighting style.

Consider landscape photography and the numerous folks who go and attempt to replicate well-known landscape images at famous locations. How many shots have we seen that attempt to replicate Ansel Adams’ Tunnel View over the past fifty years? How many Instagrammers are photographing the Empire State Building as seen through the Manhattan Bridge in the Dumbo area of New York? These shots are a bit cliché and yet we want to capture our own take on them. There’s something to be said for folks using an iconic shot as inspiration and then making their own version of the image. Each new creation is slightly different than the previous ones, sometimes in subtle ways. Are these images “worse” because they’re not identical? Not at all. If a replica image would be perfect, these images are often more interesting because of their imperfections.

Computers are great at making copies, but to replicate the beauty in imperfection we must consider whether artificial intelligence can be used to introduce imperfections and whether those computer-generated imperfections will be as pleasing to our human eyes as the original work created by people. If we look at an effort that creates tens of thousands of fake stock photos of people, those variations are created by introducing differences (are they imperfections?) into the algorithms.

Can it ever truly be an imperfection if it’s the result of an algorithm?

As AI-generated artwork becomes more common, I suspect that we’ll see human-created work valued for what it is. It will be valued for the effort involved, the time spent by its makers to learn and hone their craft, and yes, it will even be valued for the inadvertent human-caused “imperfections” in the artwork that make it unique.

Correct and/or Interesting

As photographers, we try to get better at our craft. We strive to make images that have better lighting. We attempt to pose our subjects in ways that are more appropriate for the subject. We often look to our computers and software (and increasingly to artificial intelligence features) to help us make things more "correct"... for however we define correct…

But we should also consider that photography includes both a technical and a creative component.

How often have we encountered someone who makes technically correct, yet rather boring photographs? That was certainly where *I* was at as I started my journey into serious photography.

As we look at how software will push our photography forward, we must also ensure that we don't lean on software that will push us into the corner of making more correct, yet more boring, images.

AI Will Replace the Menial Tasks

Considering the examples of automation in business cited above (grocery checkout, warehouse package assembly, simple transit shuttles), we see that the types of work that have been automated are menial tasks not requiring deep thought or a lot of applied brainwork. While there will always be particular exceptional situations, most of this work is pretty routine. Of course, as our world evolves and technology changes, the definition of a “menial task” changes as well.

In general, as artificial intelligence enables new and more advanced automation, we see the automation replacing the *doing* as opposed to replacing the *creating*.

Consider the Photographic Routine

When considering how artificial intelligence and other automation will affect photography, it's the routine menial tasks that are going to be replaced by computers. Earlier I mentioned the Square Photo Studio, a service being offered to create standard, cookie-cutter product photography images following a pattern so familiar that Square has been able to build an automated system to operate the camera to create what is essentially the same photograph, with different products swapped in front of the lens. This is a situation where the photographer's brainwork doesn't go into the capture of the image, but rather into the lighting. Once Square figured out the lighting scenario one time, they can offer this service at a lower price than it would cost a business to hire a local product photographer.

In another example of routine photographic work, imagine a group portrait... perhaps everyone on a sports team. We’ll have a couple of dozen individuals. Where does a photographer need to apply their thinking in this scenario? In all likelihood, it's in the posing of the group. You're going to deal with folks of different heights and body shapes. You'll want them posed in a way that nobody's face is obstructed. You'll want folks arranged such that the overall "shape" of the photograph is pleasing to the eye. This is what's going to separate a great pro photographer from a hobbyist making a snapshot. After the subjects have been captured (in several photos), the photographer would typically go back to their studio for some editing work. A fair amount of this editing work falls into the realm of things that can now be handled by smart software using artificial intelligence. Someone who blinked? Identify the blinking ones and automatically replace those eyes with open ones from other images captured. Who's smiling? Who's not? Again, using several images as source material, the software can identify the "best" looking version of each person for the final image. Even if the software isn't perfect, it can do a significant portion of this work, freeing

up time for the photographer.

AI will take care of more of the doing, while as photographers we will focus on the creative angle. To reference a specific example, recently I made business headshots for six individuals who work for an insurance firm downtown in my city. It was a fairly routine gig. I'd scouted the building for an ideal location, coordinated a time to show up, set up a softbox and a reflector on light stands, and spent a few minutes with each person to make their portrait.

Can artificial intelligence replace me? I doubt it… at least not for a long time.

Can artificial intelligence help with this job? It can. It can help ensure that my exposure is consistent between images. It can help take care of potential blinks or blemishes with my subjects. It can help with post-processing, image selection, and retouching.

Could it work with the client to understand their needs? Can it scout the location? Of course not.

AI can be a big help for the *doing* but not for the *creating*.

What's the Brain Cost?

Most conversations around how artificial intelligence will affect the human workforce are centered around tasks or even entire careers that could go away in the future. Heck, I used that premise as the title for this chapter. But there's another factor to consider: if we automate all of the simple or menial tasks, what's left? How will a workforce react to the changing types of work that make up their altered set of duties?

As the lower-skilled tasks become automated and removed from our list of concerns, we eliminate a lot of work that typically didn't require much (if any) decision-making. The remaining works will be of a higher level that requires more brainpower. If you eliminate the simple tasks, only complicated tasks remain.

The work gets harder.

We know the human brain typically has a finite capacity for making difficult decisions and there is a limit to the amount of deep work we can do in a day. How does the automation of the simpler tasks affect our overall work capacity and our mental health? Decision-making is cognitively draining, but when this draining work is part of a greater set of tasks we perform throughout the day the drain doesn't become excessive or incapacitating. But if our workload shifts to be much more intense and draining, the result could be that the photography work we once enjoyed will become less enjoyable or even downright unpleasant.

Even if we didn't consciously enjoy some of the more menial parts of our photography, they were part of our overall workload and how we felt about being a photographer. As those easier tasks disappear, how will we feel about photographic work becoming more mentally taxing? And if the work gets harder, will the photographic market shift to place a higher value on that work?

These are questions that will be answered in the next ten years...

And Then…

With the rapid pace of developments in technology, the topics covered in these pages have likely evolved since publication. This is a good thing. Through improvements both to the technology but also to our applications of it, photography can move forward to help us all make more impactful images while also recognizing that we have a responsibility to our shared society to use that technology in equitable ways.

For updates on the material in this book, information about the latest developments in photography technology, and insight into what's coming next, get the latest info at next.techphotoguy.com

If you found this book interesting, I'd love for you to spread the word and leave a review on Amazon.com - your support is appreciated for my work!