**Matt Powell:** 0:00

We have 3.5 million cameras globally. About 56 countries were deployed in. We have large amounts of data sets that we can pull from. It's not like we have to go to a third party and purchase a thousand images of stairwells. If you look at something like a fight detector or running detector, those are going to be 99% the data sets that you've used. That is the skeletal model. It's large motion that it's looking for. The general AI spectrum is that people think that it can do too much instead of really narrowing the focus.

**Craig Smith:** 0:33

AI might be the most important new computer technology ever. It's storming every industry and literally billions of dollars are being invested, so buckle up. The problem is that AI needs a lot of speed and processing power. So how do you compete without cost spiraling out of control? It's time to upgrade to the next generation of the cloud Oracle Cloud Infrastructure, or OCI. Oci is a single platform for your infrastructure, database application development and AI needs. Oci has 4 to 8 times the bandwidth of other clouds, offers one consistent price instead of variable regional pricing and, of course, nobody does data better than Oracle. So now you can train your AI models at twice the speed and less than half the cost of other clouds. If you want to do more and spend less, like Uber 8x8 and Databricks Mosaic, take a free test drive of OCI at oraclecom slash IonAI. Hi, I'm Craig Smith and this is IonAI. Today I speak with Matt Powell, managing director for North America of Intelligent Security Systems, or ISS. We ventured into the intricate world of video analytics and its expansive application in urban planning, safety, transportation and beyond. Matt illuminated how ISS leverages a blend of advanced computer vision models and sophisticated analytics to extract valuable data from video footage, from tracking pedestrian behavior, to enhancing city infrastructure and ensuring public safety. The discussion underscored the transformative potential of AI and shaping smart cities as we delved into the technicalities of facial recognition and the granularity of video analysis. The conversation revealed the burgeoning capabilities and the increasingly accessibility of AI in the public and private sectors. Why don't you introduce yourself? And then I'll ask some questions about ISS Sure.

**Matt Powell:** 2:47

So my name is Matt Powell and I'm our managing director for North America. So ISS, we are based out of Woodbridge, new Jersey. We're about 27 years old, 18 offices around the globe, and so we got our start back in 1996 as an industrial analytics company, which means that at the time we were just doing very basic line counting in a chocolate factory over in Eastern Europe and our founder, who's still involved with the company today, came up he said you know, if I had more footage I could do more training on analytics. And so he built his own NVR back in 1996. They weren't exactly in high demand over in Eastern Europe. So we released our first facial analytics in 2003, our first LPR vehicle analytics in 2004. And since that time we've gathered we should close. I think we closed last year with about 30 patents on different algorithms and video analytics systems. And so you know, I got my start in the industry Gosh more than 20 years ago. I worked for a company that built cameras and then worked in the integration side where we would plug all these cameras and analytics up. And you know, when I saw what ISS was doing? Globally there's about 3.5 million cameras running our analytics somewhere in the world right now over about 300,000 sites. When I saw what the company was doing, being on the integration side, I was like how has nobody really heard of this in North America? You know real big throughout the rest of the globe. So I came over, helped out with the reorganization that was going on to kind of gear up. Even though we've been based in North America for more than 20 years it's always been the home of our intellectual property I came over and kind of rebuilt the engineering group, rebuilt the sales group and you know we've been off to the races ever since, yeah, and this the video analytics.

**Craig Smith:** 5:01

Are the video cameras streaming to a central server where the analytics are done, or is it on premise? How does that work?

**Matt Powell:** 5:14

So there's two ways to do it. So A, we do it on premise. So you will have an appliance on premise, so you will have a server, typically, and then you'll have a GPU build out of, you know, intel or Nvidia cores and they will process everything on site and then from there up into the cloud. That's where you get a lot of. You know, your interactive user interface with your reports and your dashboards and your notifications and so forth. It's very difficult to run. You know, when you look at very basic analytics that come through a camera, if you're just running a basic line analytic, a lot of people can kind of. You know, nothing really exists in the cloud. You know it's not the upload and the download. That's really the challenge. It's where are you going to do the processing and then how are you actually going to, you know, send that back up to generate a notification based off of what it is so less complex you can do in the cloud when you get to handrail holding and telling if people are holding handrails as they're walking up and down stairwells. That's not something that you can do in the cloud. You're typically going to need to be on premise, but then you're able to take all of your notifications, your alerts, your dashboards, your reports, all of the data that comes out of that. You can put it into the cloud.

**Craig Smith:** 6:43

Yeah, and the detection, for example, with handrails. Why can't you do that in the cloud? Because of latency or what so?

**Matt Powell:** 6:54

there's kind of three places that this happens. You have the camera, so the camera has, you know its processor. So some cameras will actually perform some analytic function. At the edge in the camera they will also provide metadata back. So if you're looking at a crowd, the camera may provide metadata back to tell us to blur all of the faces in the crowd. So you have at the edge with the camera and then you have your appliance. So it's a little bit different than a server, because a server you know you can run analytics on a basic laptop or a desktop. But to do more accurate when you're talking about stuff with 99% accuracy for behavior, you're going to have to have some type of GPU processing power. So most modern cameras only have between 10% and 15% of the GPU capability to actually perform high-level accuracy analytics. So they can do a basic line. But once you you know, once you start getting beyond that, you need some type of appliance. Where the latency happens is in the upload and the download. So let's say, for example, we have a client right now and they have people who are jumping over turnstiles to enter into a facility instead of swiping card and then going through the turnstile, they jump the turnstile. So they want to know when that happens. It is extremely, you know, heavy processing power-wise to constantly be running that on site. So what they want is that on site. They want a notification when it happens. They want to tag that, then they send that to us. We process on site at our headquarters in New Jersey and then we send them a report on it. So to do it real time is where the challenge comes in. The latency you know, to get real-time notifications of what is happening if somebody you know has a weapon, or if somebody, like I said, is holding a handrail, if you're on a construction site and somebody is not wearing their fall harness and they're climbing up in the building, that latency in time when you notification it's just too much to be able to do that type of stuff in the cloud.

**Craig Smith:** 9:13

Yeah, that's interesting.

**Matt Powell:** 9:14

So those are the types of things that typically you're going to send off and then you get the report back. Versus being on-prem, you get the notification immediately. Then you go into the cloud and that's where you can start getting you know your reports and your dashboards.

**Craig Smith:** 9:30

Yeah, and so this, and just to clarify clarification when you said a line, what line are you referring to?

**Matt Powell:** 9:45

What I mean is a basic analytic line, so a basic pixel line that you're going to put. You know, you're going to go into the camera and you're going to set it up to where you have a basic pixel line and anything that crosses that pixel line. You want a notification, so it's a kind of line crossing. That's the most simplistic way to put it, I guess. So you know you're just drawing a basic line in the software and then anything that crosses over it you want a notification of when people cross over it Versus a skeletal model holding a handrail where you've had to teach it what a handrail is and then you've had to teach it. You know a skeletal model in relationship to the handrail and then is it touching. That's a completely different level of processing power that the GPUs on the cameras themselves just they're not at that point yet where they can process at that level.

**Craig Smith:** 10:36

Right, and so the models, the analytic models, the AI models, sit on the server on premise, the sort of first line of analysis, and that then can send an alert at very low latency to some control center on premise, and then it uploads the data to the cloud and sometime later you get a report and I would imagine that it aggregates reports from various different cameras and that sort of thing.

**Matt Powell:** 11:18

You've got it, you've nailed it, you know that's exactly how it functions. So when you're trying to get to that level of notification, it's going to have to be on premises so that you can. You know a lot of people they think in terms of the cloud. Recording takes place in the cloud and then you can take that recording and then you can run analytics on it and then reupload dashboards or reupload your metadata or your data into your dashboards and your interface. But when you think in terms of how advanced do you need I mean, the average city in the US has what I think it's six cameras per thousand people and so if you are, you know people often think of video and they see a camera up there the amount of data that comes through. If your analytics are extremely highly accurate, you're looking at a second set of eyes for whatever task that you've trained it on. So if we've trained it on very specific tasks that we want it to detect in the environment, it's like having a second set of eyes. Well, if you think of the cloud, it's almost like you know you're viewing something through miles and miles and miles away. You know there's a long bit of latency to tell somebody hey, I just saw this happen versus if you're right up on top of what you're trying to look at, then you're going to get much faster in terms of you know your reaction time to whatever that task is when you get the notification.

**Craig Smith:** 13:06

Yeah, and what kind of models are you using?

**Matt Powell:** 13:10

So that we're speaking the same language, so that you know we're speaking the same language. When you're talking about models, are you talking about, like, how we're training them? Are you talking about, like, the modules out there that people are actually able to use in terms of what data they want to learn about their environment?

**Craig Smith:** 13:30

Yeah, I'm talking about the computer vision. Presumably that's what they are computer vision models that are analyzing the live video, and is this supervised learning? Yeah, I'm just that's what I'm interested in, you know.

**Matt Powell:** 13:54

Okay. So with the software and on that software you've got that model. So that model is a module and it's looking for something very specific. So if you have cameras in a stairwell or cameras on a construction site or wherever it may be, then it is looking. The model is looking specifically for humans going up and down stairs with a handrail and whether they're touching the handrail or not. So it becomes preloaded with that. So, like you said, it's pretty light because it's very task-based. So we have about 50 different tasks that we look for and then we can combine them into different packages so you can look for multiple things with one camera angle. But the model comes pre-trained, it's already there on site. You set it up, you tell it. Okay, let me make sure the camera is positioned right Now I'm going to make sure. You know you almost you have to go in and you have to set the height of the individual that you're looking for in terms of perception, so it knows exactly where people are going to be in the scene. So this is a highly specialized one. For example, you know hand washing same thing. The model comes. The model, the skeletal model, has been trained so it knows what these different things are. So when you look at like the base model here with the convolutional neural networks that we often talk about, so for us you've got it knows what a human is and then it looks for the behavior, so based on the skeletal representation, so it's trained on the data sets that we have. So the data sets are going to be a large number of different stairwells with different you know lighting, different people, different numbers of people that are in the scene, people wearing different clothing, whatever it may be, and so we're just going to keep training it on that. And the other thing that's interesting is with that skeletal model is that we can create rules using natural language because of that model. So we can say you know, notify me if somebody's not holding the handrail, and it's going to do that notification. So all that comes into us in a software module, and that software module is what is loaded onto your appliance that arrives. So when you turn it on, software pops up. You bring your camera feed in. We you know either someone who's trained on the software or one of our people remote in. They make sure that the scene is right, hit, play. It starts looking for what that task is. So it's going to start go ahead and running it. So, like you said, it's light in terms of you know it's very, very specific to whatever it's looking for. And then you've got that GPU, horsepower, processing power there in order to process on site and start sending the notifications.

**Craig Smith:** 16:49

Yeah, how general are these different modules that are trained on different tasks, like the handrail Is that? Can that be deployed in any setting, or do you have? Does a client have to provide you with data to fine tune the model so that it's familiar with the client's environment?

**Matt Powell:** 17:24

It's about 90% out of the box, 90% of the scenarios that it's going to see. You know we have when you have 3.5 million cameras globally. About 56 countries were deployed in we have large amounts of data sets that we can pull from. So it's not like we have to go to a third party and purchase hey, you know, can you send us a number of different. You know, can you sell us you know a thousand images of stairwells. If we're going to build something like this, then we're going to look out globally, we're going to talk to everybody around in our 18 offices, we're going to pull all those data sets back in and then we're going to start training and then, you know, depending on the situation, we may use data augmentation or GANs or something like that if we've got some gaps. But typically, with that number of cameras deployed worldwide, we can reach out. Our engineering and R&D group can and they can pull massive amounts of data sets that are going to replicate about 90% of what you're going to see, because we're going to have nighttime, poor visibility, you know, like I said, different numbers of people, different signage, different frame rates and so forth. Then, once it comes back, you know for the behavior that it's looking for, accuracy-wise, depending on the neural network training time, you know you can get it up to. You know, 95% to 99% accurate In terms of generality. You've got a staircase in your home and or in your business and you want to, you know, deploy a camera in there. Probably about 90% of the time on that type of analytic, it's going to, you know, turn on, recognize everything that's going on. You have somebody walk up and down so that you level set it in terms of perception, so it knows how big people are going to be in relation to the objects and it's good If you look at other. So that's a highly specialized one. If you look at something like a fight detector or something like that, or running detector, those are going to be 99% because you've the data sets that you've used. You know that is. You know the skeletal model. It's you know kind of large motion that it's looking for. It's had massive amounts of training on what human running is. It knows what human is. It knows the difference between a human walking and running and you know, or a fight, you know it's been trained on what fight mechanics are. So it's going to go ahead and in a general scene it's going to know that very quickly. So you know being able to deploy just to a general environment. This is why some companies there was a story that came out recently about a company that they had one. It was a person down, you know, a fall detection analytic and they deployed it. Company bought it. They, company, you know this startup remote it in. Hey, we need your employee to lay down on the ground in front of it. Okay, we need them to roll to the left now, roll to the right now, you know, kind of sit up, kind of do this, kind of do that, all these different things for you know a couple of hours to train it. The next week a guy had a heart attack and he fell down in front of the camera in front of the elevator, wasn't detected, and so somebody walked out and then they found the guy and they went back and they said what happened and the answer was well, he didn't really fall in the way that we trained it when we were doing the training, and so this is pretty common now in the video space when you look at the number of cameras that are everywhere and then you know. So that is a large opportunity for companies to get into with the accessibility of you know, you know analytic training. The challenge is getting large amounts of data sets and either you have to have a lot of money to go and purchase them or you have to have a large amount of deployments globally so that you can get more and more data. There are programs out there where people will go and they will purchase you know, hey, I need 180 images of a road because I'm trying to train the analytic for a very specific purpose with vehicles. But to get highly accurate to where you don't have to have somebody roll around on the floor for two hours when you purchase it, you have to have that large amount of data set. So that's why ours, you know, works pretty much. You know, when you plug it up on, you know something that's detecting what we call like our tracking kit. It's going to just be 99% accurate for whatever it's looking for, because so much neural network training time has gone into it over the years.

**Craig Smith:** 22:16

Yeah, and you said that you have you look for various tasks. How many different tasks do you currently cover? If that's the right language.

**Matt Powell:** 22:34

Module wise, we have about 50. So those modules do all kinds of things. So if you look at, for example, numbers in the environment, so we can track numbers in the environment pretty easily Whenever numbers are moving through the environment, whether it's a barcode or whether so, if it's a barcode on a package at a farm, or if it's numbers that are written on the side of a vehicle at a car auction, or whether it's cargo containers that are coming off of a ship and it's the numbers that are on that right, so that you basically are taking that number and number and letter recognition in the environment and now you're creating different tasks out of it. So the task in a lot of ways. How many tasks do we have? It's almost up to the imagination. What do you want to know about your environment? Because we can apply it in a lot of ways. So let me give you an example. We know if somebody is sitting or sitting. So that is a behavior that we can detect. The task may be that a business wants to know how long. So let's say that you have one of those cooperative work spaces. How long are people sitting at a desk and when do they get up? So there's a task. How long is their laptop open? How long are they working with a phone? So we know what a laptop is, so we're able to say, okay, they've been on the laptop, they had the laptop open, it was existent and open for a period of time. That's a task. They were holding a phone for a period of time. That's a task. So it's really up to the creativity that you can start to look at and go, okay, I have a second set of eyes that this knows what behavior it is, it knows what people are, it knows all these different things. What do I want to know with these cameras that are in the environment? What data do I want to get to make better decisions or whatever it may be? And so I'm going to just apply these analytics to that and then we go in and we kind of adapt around that. So the tasks you know module-wise, there's about 50 of them, but task-wise it's really limited by the imagination and obviously, how much money people have to develop certain capabilities. But you can apply this in a lot of different directions as long as we have a kind of a clear scope of what you're looking for and you say, okay, I'm looking for this. If we don't have it in existence, then we go out, we try to pull the data sets, bring it back, start to train it. You know, we'll set up a little studio at our manufacturing facility. We'll reenact seams in there number of different things that you can do if you need to. You know, apply these analytics to detect something and give you data.

**Craig Smith:** 25:48

With everything that's happening in multimodal models and computer vision. How is the tech changing? Is it changing for you guys or in the marketplace?

**Matt Powell:** 26:11

There's two ways that it's changing to a certain extent. So it's becoming more and more accessible, and you're seeing more and more companies able to do more things that they used to not be able to. And the other way that it's changing things is that people are. You know, the amount of bad experiences that happen with it are making people question it, and so the changes that we see is that, as it becomes more accessible, more and more companies get into this arena and they're either licensing somebody else's analytics or they're generating their own, or they are getting license-free analytics from a company and then developing a bit of their own interface and then reselling it. There's a lot of that going on, and so that you know, it puts it out and it makes it more popular, so more and more people look at it and people become more and more familiar with it. But the other side is that people try, instead of making it task-based. When we talk about AI being task-based, you know it's just trying to replicate a human action with a machine or an algorithm they're trying to do too much with it, and so in doing too much with it, you end up with an area of the market that is been over-marketed to, and then people have been let down by it, and so I think that's kind of all over. The general AI spectrum is that people think that it can do too much instead of really narrowing the focus. So what we've seen is that the consumer is becoming more educated in asking the right questions so that they kind of narrow it in as more and more of this technology gets out there. So I think it's a positive thing, change-wise, the other thing that we've seen is that the processing power is making things possible that weren't possible. So we're fortunate we're not a start-up. We've been around this for a long time. So when you start looking at vision transformers and so forth, they're extremely heavy to run and the processing power is something that a lot of companies have challenges with. But us, we have partners, so they assist us with computational capabilities. So we have partners like Intel and Nvidia, and we might train on Nvidia, port it over to OpenVINO and bring it into Intel, not to go too much in how the sausage is made. That was really difficult to do years back now, because when you see these companies coming out with new chips and you see these companies coming out with new capabilities, I think we're starting to become more and more accessible because the computational power is there. So we're able to dig more and more into things that we couldn't do before, because the computational power is increasing every year so it allows us to do more than we could. So it's a rapidly changing space and it's exciting to see it because consumers are getting more educated. They're asking the right questions. And the computational power when I talked about earlier, the cameras can only do between 10 and 15% of what the GPU horsepower or computing power is to be able to run these at the edge. The estimate was in five years they'd be about 75% capable, but now people are revising that back and saying possibly three. So what are you going to be? It just leaps and bounds every year, and so what happens when you get to that part? It's not only that. But now you're looking at when you talk about change, you're looking at the security integration world is a multi-billions dollar industry. Now, when you have companies coming into it from outside of the security world and they're able to work with cameras and they're able to work with video, that is coming in and you've got this massive retraining that's happening and you've got tens of thousands of people that are trained on deploying a server, that what happens when the server starts to disappear and it's all out at the edge. These are questions that every industry that is working with AI is having to grapple with constantly, and the video portion of this is making this data more accessible. This security used to be a very closed world and it's been pride, open by the accessibility to these types of analytic and AI capabilities, that now more and more people are getting into it and offering more and more stuff. We see startups that come in and they do absolutely. It's very task based, but it's absolutely incredible stuff that five years ago it didn't exist and now today they're taking over quick serve restaurants very rapidly with just one little thing that they do with one camera, but they provide massive amounts of data. That wasn't possible five years ago. Change is very quick when it comes to this side of the AI house.

**Craig Smith:** 31:59

What are the largest categories of deployments? Is it security? Is it safety? Is it in factory settings or hospital settings? Can you break down your customer base?

**Matt Powell:** 32:22

The largest that we see is it depends on complexity. Mexico City we have Mexico City that is 60,000 cameras. I want to say it is so when I was talking earlier about running and knowing if people are running or fighting or whatever it may be, I think that has 10,000 of those modules. It can track people wherever they go. It's got facial I think there's 1,000 license plate recognition, 1,000 facial recognition and then you've got five command centers. You've got a false alarm AI when people are calling in, that if they're doing false alarms, you can train it so that it recognizes false alarm calls that come in. Then you've got 13,000 kiosks around the city that all have two-way video so that you can push a button and they have a 24-hour medical center where people, if they're on the street, they can push the button and it pops up somebody. There's 30 full-time people, 24-7 in this medical center that sit in front of a camera and can have a two-way conversation with you. In 13,000 locations around Mexico City. If you have a medical event, that is a gigantic and that is a hybrid cloud system that we run down there. Then all of a sudden you back off and you've got some factories that they're even more complex. So they may only have 100 cameras, but they're running 10 to 15 of these modules on each one of them detecting all kinds of safety things that are happening. So their notifications are in the complexity of the data that they're putting off. That they go through and they look for safety related is a very different thing. So we see cities a lot, we do a lot of sports venues, so we have our occupancy counting and how many people are coming in. We have a bunch of vehicle things. We have our under vehicle system that scans the bottom of a vehicle and then uses AI to understand if there's any type of modifications to the undercarriage of the vehicle and sends an automatic notification that it's trained with grenades and bombs and it has a magnetometer to detect changes in metal alloy. So if you look at like FIFA, we did 40 of those units, but the amount of data that they're putting off is massive. So it really just depends. If I had to rank it, I would say cities are the largest commercial settings where people want data on what's going on in their environment. Manufacturing facilities is extremely. Those become pretty large. Hospitals are very big because we're transitioning into a world of virtual care and so virtual care allows us to pop up a camera in every single room and then know if somebody has fallen out of the bed, know if somebody, if you have patient elopement, where somebody is walking out of an area that they're not supposed to, so you go into a hospital or hospital chain and now you've got 500 rooms in the hospital. You multiply that times 50 hospitals in a hospital system. You're looking at large amounts of cameras, large amounts of data, which is why being on-prem is so important to be able to process all of that on-prem Trying to do. You know this isn't for four cameras. You know if you're doing you know two 300 cameras at a time, you're going to end up with a large amount of data that you need to process on-site to get immediate notifications of what's going on. So really just depends. You go from 60,000 in a city to where you go to. You know 100 cameras in a manufacturing facility or you know a farming facility, but the amount of data that it puts off, that they go back through and they audit, is incredibly impressive with what it provides you in terms of data.

**Craig Smith:** 37:00

And are these systems getting cheaper? I mean, certainly the hardware is getting cheaper. Models are proliferating. Is it getting to the point where consumers are going to be able to deploy this kind of video analytics?

**Matt Powell:** 37:22

It's a good question that it is definitely getting cheaper in terms of how many things that you see out there that promise stuff, and it's getting less expensive as well. But I would say coming to the consumer. So for businesses it's getting a lot less expensive. To do some of this stuff years ago was cost prohibitive, recording and video management being completely different. But nowadays, depending you can go, for example, through a company like us and you're able to pick out the modules that you want. You can have your integrator build your hardware, or we can build the hardware or you can build the hardware yourself at this point and then you can go. We have tons of business owners. They build the hardware themselves, install it, they pay for their service and maintenance over time, so they get all their updates and so forth and then they're off and running and they're at the same level as when you see on TV and they AI for your business and you think it's for the big guys. It's not. It is completely democratized at this point, very consumer friendly. You can get access to it for business. It's a little bit different. A lot of companies will not sell to residential companies like us. We won't do residential unless you're talking about the governor's house being a residence or the president's house being a residence, and the reason is because the technology is typically going to be above your average consumer to be able to scale it the amount of customer service challenges that you would get out of selling into houses all over the place. On high end analytics, right now for residential you can get basic motion analytics and basic things like that. But if you're talking about higher end types of analytics, I don't know if it will ever get to that point where people can do it. There are residential companies that are providing more and more video capability, but the amount of, like I said, scale to release high end analytics into the residential consumer world, the scale is just not there for higher end stuff but for lower end, basic security, basic things in your house. It's there today that people can find.

**Craig Smith:** 40:04

Yeah, I mean I see people talking about facial recognition locks and things like that, and I've spoken to people about elder care and this idea of outfitting homes with multiple sensors, including cameras, that then learn the patterns of behavior of the person who's living there and then can alert caregivers or family if there's a change in behavior, whether it's a fall or just a change in behavior that might indicate increasing dementia or something. Are you saying that those applications still don't reach the level of sophistication of what you guys are doing and that's a cost issue, or is it coming?

**Matt Powell:** 41:16

So facial recognition for door locks is very different. So let me give you this example If you put on sunglasses, you know so you know mobile phones today that use facial recognition to access them. If you put on a hat and sunglasses, they start having problems. To get to the point where when people wear sunglasses, you are still detecting 40 points on the face so that if they put on a mask or they put on sunglasses, they can still access an area is a different level of facial capability because it's moved beyond face to pattern recognition and patterns on the face. And they just happen to be patterns on the face. They could be patterns in a fingerprint or patterns in your ear, because ears are just as unique as fingerprints. So the basics of having your phone turn in your face left and right in order to access is very different than having facial to the point where you can use it for multi-factor authentication and is a biometric to get into a highly secured area. So you're going to have that area of the market start to consumerize, where people will get that because they will be willing to deal with a lower level of accuracy than the commercial world is, and so there are companies that will move towards that and then you'll have larger companies like Amazon and so forth that will start to do that. There have been kind of when you look at elder care, starting with fall management, where people have started to get a lot of startups that got a lot of money pumped into them For that. They sell it into the hospitals. The hospitals start to purchase it, the hardware starts to decrease and the cost of the analytics starts to decrease and then it starts to be something that you can begin to put into the consumer's home. So it's the same thing when you look at the analog camera and the IP camera. And when the casinos started purchasing tens of thousands of IP cameras, it made all the components and everything else to the cost point where it didn't make sense to make analog cameras anymore. So everybody started making IP cameras and it started to be larger adoption. So you kind of have to have this ramp up time, obviously, where people are going to adopt certain things. The question will be is there going to be that market where someone has all the mechanics to be able to go and charge thousands of individual homes and monitor for these things? The technology exists right now to put it in home and tell if somebody's fallen down. You can train certain behaviors. We have stuff for farms where we know if cows are going into labor, so you can train it on certain behaviors, because pigs and cows have a certain behavior when they're starting to go into labor and so if you own a farm you kind of want to know that if it's happening in the middle of the night and get a notification so you can go out and you can take care of it. That stuff already exists. So elder care type of technology already exists. The question is going to be at what point does the industry grow to where companies can support the ability to deploy it? My personal opinion is that I think you will probably see companies that are out there, the gigantic organizations that put cameras everywhere that they possibly can to gather data, and they will start to find that as a means for gathering the data that is out there. Because it's very expensive to deploy this stuff into houses. You see it in hospitals and I know the cost on a hospital to be able to do it. I can't imagine how you would be able to do it in the next couple of years in a residential home without having to add on a bunch of different stuff that really invades privacy so that you can gather massive amounts of metadata and resell it. That being said, like I talked about earlier, everybody thought that the edge cameras would take five years to get to the point where they would have 60 to 75%. Now people think it'll be about three. So I think this stuff is moving so rapidly that we could see that just speed up. It's eventually going to happen in individual homes to be able to detect this if people want. That Question is the scale and cost for the companies to own that type of intellectual property. It'll get here, whether it happens in the next three years, five years or 10 years is going to be the question.

**Craig Smith:** 46:26

Yeah, do you guys use your system at all for online monitoring? Facebook has invested a lot in scanning live video feeds for signs of violence and pornography or something.

**Matt Powell:** 46:48

We do have a managed services group that will process video for people, so we do have that. It is not something that we generally do. It is something that we have a managed services group that has the capability to do things like that, where you could scan live video because that's all. Video is right, you got a camera out there and it just happens to go back through your server versus. But online monitoring is something. It's not something that we majorly do at this point. Our managed services group is they can take video and they can start to do those things and do run the analytics on them and then provide data off of it, but it's nothing that we've really looked at in terms of the scale of people. Like Facebook, they have the infrastructure to do those types of things is different than the infrastructure for a video intelligence company like us. That goes out and we might do cities or we might do manufacturing plants. When you're looking at that, it's just a different level of infrastructure. A lot of the technology is the same Like you said, monitor for violence happening in online videos or videos that are live streaming but the infrastructure is very different to be able to accomplish that.

**Craig Smith:** 48:18

When a city comes to you guys, who generally is it? Is it the police department or the mayor's office? Or who is buying and implementing this stuff?

**Matt Powell:** 48:34

So there's two areas that come to us. A lot of who comes to us is the traffic engineering and city planning, so we have a large capability to do pedestrian analytics and then you're able to extrapolate data from there. So let's say that, for example, you have a lot of reports of near misses that are happening in a mid-block area, so you'll go out there or you have unfortunate fatality. So they want to know when people are crossing, instead of at the crosswalks. They want to know when they're crossing and then they want to get a count of that. When's it happening? And then they want to look for why is it happening? Is it related to the bus schedule? And there's an apartment complex and the transit bench happens to be right across from the apartment complex and instead of walking to the crosswalk, they walk right across the street to get to it. So they'll want to put up cameras or use live feeds and install a module that's going to look for pedestrians doing stuff that is anomaly and look for them any time that they're outside of it. We have a system called the SOFIT that will actually illuminate people. It's an intelligent lighting solution, so as they cross the crosswalk, we use AI to drive the lighting so it illuminates the individual as they go across the crosswalk. So a lot of times the license plate recognition side, cities call. But there's a lot of companies out there that do license plate recognition and they use their automatic license plate readers and then they're connected into different databases. So we get calls for that from cities. We've got a lot of interest right now in that under vehicle system for a lot of the correctional facilities and a lot of areas where they're walking around with mirrors to look under vehicles. But most of the calls that we get when it comes to a city are about kind of building that smart infrastructure to where we're able to connect into traffic lights, run traffic lights. We do connected vehicles so we are connected straight into a vehicle and tell them when a pedestrian is going to be in a crosswalk that they're driving up to, to counting pedestrian actions, looking for traffic backups. All of these different things kind of go into that intelligent transportation systems area of the cities and city planning. So we deal with a lot of traffic engineering departments that are trying to understand that and kind of build that backbone for themselves that you can get more of that connected vehicles and vehicle to everything the V2X, cvx systems. That's the future when you really look at it. Being able to have all of these things connect in the cities is going to. We see a little bit of it now, but five years from now, the things that we see coming is absolutely incredible of what you're gonna be able to do with AI to rapidly speed up all elements of city planning and then understand reducing pedestrian crashes and fatalities and injuries. That type of stuff is amazing. That's something that we do globally. It's kind of been our DNA, because it all goes back to that skeletal model right Going back to 2004 with those first LPR analytics. We've been working with vehicles for 19 years now in terms of neural network training and training our algorithms. Now you take that vehicle and we know what the vehicle is. We can do the make, the model, the classification. We know a vehicle in its environment. We have the skeletal model. So now we've got that human that's out there. So we start to combine these things. Where you have this environment, that is the roadway, that is the sidewalk, and you can start to make these systems start to interact with each other to where they're all working together at once and that's where all those modules start to come together and stack on top of each other. So now you're getting all of that data, but you're also actively. You're able to work in the environment to change traffic lights, to illuminate pedestrians as they walk across, across walk, to be sending data so that you can understand the dynamics of what's going on and whether you need to change the way that you're gonna do your road construction down the road. So it's kind of funny how it all comes back to those very basic things, with neural network training and the convolutional neural network training that we do, and now moving into vision transformers and so forth to even get more out of these scenes that we're seeing so that we can do more in terms of the data that we can generate.

**Craig Smith:** 53:48

Yeah, how fine grained analytics or analysis can you perform, and I'm thinking in cities. You know there's been a lot of talk about weapon detection and facial recognition, these cameras that you're deploying in cities. Are they high enough resolution and are the algorithms powerful enough to do some of those things?

**Matt Powell:** 54:22

To do facial to get it started. So there's four steps to facial quote recognition. To start the first step, you need 60 pixels between the eyebrows. So as long as we can get 60 pixels between the eyebrows, we can go ahead and move towards feature extraction, liveness checking and so forth. Most two megapixel cameras will give you that, but you gotta be close to it, you know. So you can set it out to 10 feet. Most two megapixel cameras out to about 10 feet will be pretty reliable in terms of knowing that there's faces out there and go ahead and start the loop to go back and look for the pattern that is in the database that it's looking for to see if it's matching. So once it gets those 60 pixels, in terms of other things so small stuff and so forth you know I've seen some incredible things. You know speaking outside of our company, where you're seeing AI use for cancer detection in. You know slides. You know, instead of you know shipping. You know a blood sample or blood slide across the country. Now you can just, you know, put it online and then you can run AI and see what you can see. That is not stuff that we do. We're more about the environment in general, if you look at all the different things that we can detect, you've got in there. You've got like carrots and cucumbers and I asked our R&D team. I said why? And they said well, down in South America we have a lot of horse farms. We wanna know if people are giving you know. Are they giving a carrot to a horse or are they giving you know something else to the horse? So we've had to train it on what these things are. Weapons detection is very interesting. We have a study that's out where we had a transit authority that asked us to come in and do all these different detections, and one of them was, you know, using an object to strike an ATM, and the number of times that, when you first get started on this, looking at the reason that they came to us, was, is that the low accuracy analytics didn't know the difference between a baton and a phone because it was a black device in a hand, and so that is where you have to get into the behavior side. It is not enough. If you look at you know it's not enough just to do the object, you have to have the behavior behind it, brandishing. When it comes to weapons detection, you know that's one of the things that they found over time is is the person is walking with the firearm, you know. So you're detecting the firearm, but then you have to have really high accuracy analytics for that. But then you also have the behavior around it to set the context. Reason being you don't wanna false alarm factory, so people don't trust it anymore. They go well every time somebody comes out and they're holding a book that thinks it's a pistol. So two megapixel cameras will get you almost all the time to where you need to be in terms of the task. And, you know, the worst thing that can happen is is that, a you have to put up more of them, or B you just have to, you know, bring the activity that you're looking for closer to the camera the more minute the thing that you're looking for is. So if you're looking for something smaller, you're looking for hands to, you know, make movements and do things and so forth you just have to bring that closer to the camera. But about two megapixel will give you what you need, whether it's facial or behavior or weapons detection, but then you, you know at that point what are you doing with it and that's where the data sets neural network training time and the accuracy really comes into play. So you don't generate false alarms, you get good data that you can make decisions off of.

**Craig Smith:** 58:24

Is this? You said cities are one of your biggest market segments. Is that primarily overseas, or is there a lot of that in the United States?

**Matt Powell:** 58:39

Primarily overseas. We have a lot of cities that we work with and departments of transportation and so forth that we work with, but that is not. That is ITS related technology, that is, vehicle counts, pedestrian behavior, pedestrian safety, connected vehicle. You know, really applied analytics is sometimes what I like to call it. You know it's augmented intelligence because we're providing really good data, and then you know it's kind of applied analytics. How are we applying these different models that we have through the modules in order to give you data? You know we do have cities that we partner with to provide. You know, license plate recognition, missing persons. You know when it comes to facial and certain areas that may be. You know areas that you know. Unfortunately, you have a lot of. You know people that are victims of trafficking and so forth that you know you are able to connect to a missing persons database and then be able to look. But that is a minute portion compared to intelligent transportation and looking at the roadways and what you can do with roadways.

**Craig Smith:** 59:56

Yeah, wow, it's fascinating and I would guess it's a growing business and it's sort of an endless market. Is that right?

**Matt Powell:** 1:00:11

Yeah, if you don't do one thing. I think these days, if you do one thing, you either got to be real specific about that one thing that you do and hope there's a big market to grow your business. Our CEO, al, likes to say that. You know, whenever somebody says, hey, do we do this? He's like we've been around for so long. He's probably sitting in the back of a desk drawer. Just pull it out, dust it off, make sure it works with. You know we tried it at some point. So you know we do a lot, but we don't look. We look at it. As you know, we've transitioned. You know we're a video intelligence company and a data company providing you know data from you know what's in the environment, so it is a great business for us. You know, whether it's safety or whether it's operational efficiency, you know it's growing and the capabilities are growing and it is exciting to see, like we talked about earlier, the processing power change become more accessible and people being able to, you know, really deploy this to their businesses and their environments and their cities, their hospitals and so forth, to be able to make a difference.

**Craig Smith:** 1:01:19

Is that? Are there systems that are high enough resolution that can recognize a face at some great distance? Or is that still science fiction?

**Matt Powell:** 1:01:44

It is. This is one of those weird. It is completely possible, completely possible to do. It's all about resolution If you can zoom in and be able to get to. We do something interesting and people are always shocked One of the only companies that can render a 2D image into a 3D biometric. You can take a photo off social media or LinkedIn, whatever it is, and upload it into our face sex system. Then we can use that as multi-factor authentication, because we're not looking at the face, we're looking at the patterns the distance between the edge of your eyes and your nostril and the side of your mouth and the curvature of your cheeks and your chin and so forth. Like that. Those tend not to drastically change over the years. By the time it would change, you would have an updated photo of it, as gravity affects us all, and so forth. You're going to end up with another photo at some point and you upload it into the system. Being able to zoom out and be able to view someone's face is a matter of resolution when you get down to the basic science of it, if you can get the resolution from zooming and that's your biggest problem if you're using a large telescopic lens to be able to zoom in. If you can get that 60 pixels between the eyebrows, you're going to be able to start the process. To look, it's the same thing. You can take a small little image from social media. You can feed it into the image and it's going to be able to do the feature extraction and be able to understand. Okay, there's a pattern to this face. Let me work with it. It's really a matter of resolution. A lot of times these just come down to cost. People get scared of facial technologies and, oh, you're recording my face. The amount of storage you have to have to record that many faces, it's just not there. When you get to the level of what we do. It's a database verification. I hate the word facial because it doesn't really. It's just looking for patterns that exist on a face. But I can train it on the same patterns on your ear. I can train it on the same patterns for a fingerprint. It's just a pattern recognition. A lot of it comes down to cost. If you can afford that much of a zoom, telescopic lens to reach out and view somebody at distance, if you can get those 60 pixels in general, you are going to be able to start to do that feature extraction and look for the patterns on the face.

**Craig Smith:** 1:04:41

AI might be the most important new computer technology ever. It's storming every industry and literally billions of dollars are being invested, so buckle up. The problem is that AI needs a lot of speed and processing power. So how do you compete without cost spiraling out of control? It's time to upgrade to the next generation of the cloud Oracle Cloud Infrastructure, or OCI. Oci is a single platform for your infrastructure, database, application development and AI needs. Oci has four to eight times the bandwidth of other clouds, offers one consistent price instead of variable regional pricing and, of course, nobody does data better than Oracle. So now you can train your AI models at twice the speed and less than half the cost of other clouds. If you want to do more and spend less, like Uber 8x8 and Databricks Mosaic, take a free test drive of OCI at oraclecom slash IonAI. I hope you find the conversation as enlightening as I did.