Intel Corporation

Amplifying Computer Vision with the Power of Edge & Cloud

CORPORATE PARTICIPANTS

Lilian Veras
Moderator

Arsalaan Kashif
Happiest Minds – Director of Marketing

Ritesh Gupta
Happiest Minds – CTO, Product Engineering Services

Srikant Sowmyanarayanan
Happiest Minds – Senior Director, Hi-Tech

PRESENTATION

Lilian Veras

Welcome, everyone, to the Intel Network Builders webinar program. Thank you for taking the time to join us today for our presentation titled: “Amplifying Computer Vision with the Power of Edge & Cloud”.

Before we get started, I want to point out some of the features of the BrightTALK tool that may improve your experience. There's a Questions tab below your viewer. I encourage our live audience to please ask questions at any time. Our presenters will hold answering them until the end of the presentation.

Below your viewing screen, you will also find an Attachments tab with additional documentation and reference materials.

Finally, at the end of the presentation, please take the time to provide feedback using the Rating tab. We value your thoughts and we'll use the information to improve our future webinars.

Intel Network Builders webinar series takes place live twice a month, so check the channel to see what's upcoming and access our growing library of recorded content. In addition to the resources you see here from our partners, we also offer a comprehensive NFV and SDN training program through Intel Network Builders University. You can find a link to this program in the Attachments tab, as well as a link to the Intel Network Builders Newsletter.

Today we're pleased to welcome Ritesh Gupta and Srikant S. from Happiest Minds. Happiest Minds is a next generation technology services company focused on digital transformation and engineering, and also an Intel Network Builder Gold Partner. Arsalaan Kashif, Director of Marketing at Happiest Minds, will introduce our speakers today.

Arsalaan, over to you to start off, thank you.

Arsalaan Kashif

Thank you, Lilian. It’s great to be back for a second time this year on the Intel Network Builders channel on BrightTALK, and this time around, we are going to be talking about Computer Vision. And joining me for this is Ritesh Gupta, who's the CTO of the Product Engineering business unit, and Srikant Sowmyanarayanan, who heads the Hi-Tech business for us as part of the BU.

Now, what we're trying to do here is really unravel the various different layers of Computer Vision, try to understand some of the challenges, outline a few use cases for the audience, and finally get into the crux of the presentation, which is how distributed computing and the power of Edge and cloud can really go a long way in terms of amplifying the value of Computer Vision.

So before we get into it, I thought that I should start off with a brief introduction about the company because not everyone on the call might be aware of Happiest Minds. So we're a next generation digital transformation company that is focused predominantly on

disruptive technologies like IoT, AI, Blockchain, SDN. The list goes on and on, but we essentially position ourselves as a 'Born Digital. Born Agile' enterprise, and we started off in 2011, and earlier this year, in the month of August, we completed 10 years in existence. We had a pretty successful IPO last year. And in addition to our technology focus, one of the objectives in setting up the company was to create a truly differentiated culture, that was people first, and it was defined and rooted in the idea of being mindful and doing mindful. And I think that is yielded rich dividends for us, because you can see that being reflected in our Glassdoor rating, for instance, which is 4.3, and the second highest for Indian IT services companies, or for that matter, we happen to be among the top 25 Best Places to Work For as per the GPTW Survey of 2021.

So that's a little something for you just to have enough context, and if you'd like to know more about the company, you can log on to our website, which is www.happiestminds.com, or simply send us a message on LinkedIn, and one of us will be sure to get back to you.

So with that being said, let's jump right into the topic of the day, the reason why we're all here, which is Computer Vision, and I thought we should start right from the beginning, because there is a little bit of conjecture about the definition of Computer Vision itself and what it actually entails, because some people still believe that Computer Vision is just all about image analytics. And if you really take a step back and think about it, Computer Vision has been around for a few decades. It's not entirely a brand new concept, and it has gained a bit of traction in the recent past, and I think for us to really do justice to this, we should start right at the beginning, and for that I'd like to call Ritesh into the conversation at this point.

And Ritesh, if you'd be so kind, if you can just go in and set the table for us and help the audience understand what Computer Vision truly is and elaborate upon everything that it actually encompasses.

Ritesh Gupta

Sure, Arsalaan, and you rightly said there is a little bit of confusion when we talk about Computer Vision.

Now, let's start from the basics. If we have to compare Computer Vision with the human body, the best thing to compare, or the thing to compare is an eye. The eye of a human being is actually Computer Vision to a machine. Now, the way a human body, we see through our eyes and we analyze what we see in our brain, and the body takes certain voluntary and involuntary action. This whole thing if we map into a computer system is actually a Computer Vision. Computer vision actually allows a system to see things. Similarly, the way I gave an example of our body, Computer Vision allows a machine, a system to see things, analyze what it says, or what it sees, and then take certain decisions. Now, it is actually a branch of artificial intelligence, which trains a system to understand and interpret the visual world. Now, when we compare or when we say it is a branch of an artificial intelligence, artificial intelligence enables a computer system to think. But Computer Vision adds vision to it.

Now, if we clump all these things together, it is actually a technology which builds an artificial system by obtaining information from multi-dimensional data. Now, these data could be multiple things in terms of videos, images, and fused with certain text data as well. Now, that being said, it is no more just an image analytics. It is actually a vision, which takes the video, breaks it down into images, analyzes the images and takes certain decisions, and these decisions can be acted upon automatically by the system itself, or could be with human intervention.

Now, let's take a couple of examples, in terms of the autonomous cars. We all know about these autonomous cars. Tesla are into it, multiple other companies are into it, a lot of R&D is happening around that. It is equipped with multiple cameras and sensors. Now, these cameras are actually the eyes of the driver who's driving this car. It looks at certain road signs and analyzes that road sign. For example, if it sees a stop sign, it looks at it, it analyzes this, it figures it out that it's a stop sign, and the necessary actions taken onto the car, whether it slows down or whether it stops. Similarly, in a retail world, there are cameras placed in multiple places. These cameras continuously look at people walking around, tries to figure out what they are doing, what products they are buying, and if they are picking up certain products from a certain shelf, they can also keep a count of the number of products which are remaining on the shelf, and then can take necessary action if the stock is going down to send a notification to the store manager to replenish the stock out there.

But that's what Computer Vision is. It includes a lot of things which we are going to talk a little later during this session, but then as Arsalaan, you rightly said, Computer Vision is nothing new. It has been there, if we have to say, since 1960s, and was used in more of a research and academic area.

Arsalaan Kashif

Sure, thanks Ritesh. Srikant, coming to you, I know that you have a keen interest in the space and have been keeping a close eye on it, pun absolutely intended. I'm sure it comes up a lot in your conversations with customers and prospects across various different industries. If I had to put you on the spot and ask you, what do you think are some of the major contributors that are actually driving this exponential growth in Computer Vision that we've seen, especially in the last six to seven years?

Srikant Sowmyanarayanan

Sure, Arsalaan. As Ritesh mentioned, Computer Vision is not something which has been new. It has existed for quite some time and people have used basically the image analytics part of it. For example, in law enforcement, people like FBI and police have used the fingerprint analysis to basically catch criminals over there. However, the traditional Computer Vision, which was to work on images, was a very cumbersome and manual process. It used to take a lot of time. You had to capture a lot of different images from different angles, then you had to stitch these together, you had a manual intervention to mark various different specific areas, basically the annotation of the images and all of those things. And then you also needed a huge amount of compute power, so you needed very powerful servers to do the analysis on these images, to do the comparison, etc.

Now, some of the advancements in the technologies have actually enabled this process to become more easier and widespread. So availability of GPU-based hardware on the Edge, like Intel's Movidius Stick, has allowed us to put in some of the algorithms right where we are capturing the images and videos, enabling more real-time processing. Secondly, again, there is a huge amount of data which is getting generated and this data is coming in, due to the proliferation of multiple image capture devices like CCTVs, etc, as well as a lot of training data is getting generated, due to multiple different forums where this Computer Vision is being researched into. So availability of these large data sets has helped in training these models much better compared to the earlier processes, and also, because of the larger data sets, it also needs more automation in terms of being able to draw insights automatically from these images and videos. So that is also driving increased demand on the Computer Vision applications where the businesses are saying, can I use this for restocking my shelf, as Ritesh mentioned, or it could be on autonomous cars, and so on and so forth.

Lastly, one of the most important factors in the emergence of Computer Vision is actually the emergence of drones itself, right, and as drones have proliferated, we are already seeing more than a million drones being commercially used across the globe to do various activities, including delivery of packages, including survey of areas, like flooded areas or drought regions, law enforcement monitoring, and so on, and so forth. Now, these drones are actually pumping in so many video streams, so it is Computer Vision that is actually helping analyze these streams and enabling the multiple different business applications on top of these video streams captured by the drones.

Arsalaan Kashif

Sure, thanks, Srikant. So I agree with you that I think a lot of these factors are really contributing towards the exponential growth in Computer Vision, and it's at the doorstep of breaking into the mainstream, so to speak. But honestly, I think we're still a long way away from truly realizing the full potential of Computer Vision, and capitalizing on the benefits that it has to offer across so many different industries. And I really think whenever it comes to adopting a new technology, there are just a number of impediments related to infrastructure, security, scalability, costs. The list goes on and on, but it also remains the same, irrespective of which technology you're talking about, with a few unique elements thrown in here and there.

So Ritesh, can you talk to us maybe a little more in detail about what exactly are the roadblocks for Computer Vision, and really help our listeners gain a deeper understanding of them, so that they realize the exact roadblocks and pitfalls that they would need to avoid?

Ritesh Gupta

Sure, sure, Arsalaan. We talked about the definition of Computer Vision, and when we look at the definition of Computer Vision, it looks a little simple, but it is actually not a simple system. It details complexity at multiple steps. Now, when we talk about complexity, this could range from, let's say, building models, getting the data sets, or using the right algorithms, the detector algorithms, continuously or retraining the models to continuously improve the inferences, and definitely the infrastructure side of it.

When we talk about the infrastructure, the obvious choice for the Computer Vision is actually cloud. But then-- and the reason why it is an obvious choice is because of the flexibility a centralized management-- a centralized place where the infrastructure could be managed. But cloud could lead to other problems as well. Now, let's dig down a little bit in detail when we talk about these cloud infrastructures being used in Computer Vision applications.

The most important challenge, or most important problem, is basically the latency. What we have seen in a Computer Vision application, these are typically real-time applications. The decision needs to be taken real-time. Now, if a device, a camera, whether it is fitted on the drone, or whether it is fitted on a robot, or let's say, a retail store, if these devices start sending the video streams onto the cloud, this adds up the latency. Now, a decision cannot be real-time. It goes to the cloud, it processes, it analyzes, does the inferences, and then sends the result back to the device to take certain actions. Now, these real-time use cases, for example, there is an intrusion in somebody's house, and there is a camera which is sending the data back onto the cloud, for the inferences and everything, by the time the damage is done. So this latency actually is one of the challenges into the Computer Vision application when we use a cloud infrastructure for these application.

The other thing related similarly on these lines are the cloud infrastructure, if all devices start sending the data onto the cloud, cloud gets overwhelmed with the data that it receives, that it has to process. Now, one of the examples that we took during our Computer Vision definition is these autonomous cars. Now, let's think about Tesla. If they start-- all the autonomous cars start sending the data, the road signs that they have seen, onto the cloud, the amount of data that all the car start sending, the decision cannot be taken real-time. Cloud becomes overwhelmed, as I said, with the amount of data that it gets. Now, not just the cars, these devices could be anything. It could be a robotic arm, it could be drones, it could be robots, static camera cars, any other devices, which is capable of capturing video data. Now, apart from that, these cloud infrastructures, though, as we said, the obvious choice is cloud for the Computer Vision, but then the cost on the cloud can skyrocket if the management or the infrastructure is not managed properly, because this would continuously demand the need of higher compute, higher memory. So the cost of the infrastructure may skyrocket in this case.

The other most important challenge or problem is the data privacy. Now, when we talk about data privacy and the example we took in terms of a camera fitted in a person's house, if this data gets streamed onto the cloud, and there is a breach of this data, there is a breach in privacy. So the security brings in a bigger challenge, a bigger problem, but still there are compliances, there are regulations that are coming. In California, there are multiple legal discussions happening around the data privacy for the Computer Vision applications.

Now, apart from that, the other challenges were, which Srikant briefly touched upon, is now getting addressed is the data set. Earlier, Computer Vision needs a huge amount of data sets on which the models could be trained and that real inferences can happen. Accuracy of the inferences are, let's say, more than 90%. Now, for that we need data sets. Now, these data sets earlier was not available. Now, data sets are available, but they are not annotated and it becomes a human, or some system needs to be used, or somebody needs to sit on these data sets and annotate the images so that the models could be trained.

The other challenge that comes in my mind is the retraining part of it. We talked about initially the challenges, talked about the retraining part. Now, models once built is not a foolproof model, because the images keep changing, the model needs to retrain itself. The real-time images that get captured or videos that get captured from the devices, those images need to be used and the models need to be retrained. What it adds or what value it brings in, the model keeps updating itself based on the real-time data, so that is also very important. The infrastructure for this retraining of the models is a little early, and it's still evolving, which we call it an ML ops. Now, if we look at all the challenges together, most of the things falls onto the infra side of it. Now, that being said, and that's the reason, to scale a Computer Vision application, it is very important to distribute the load between the Edge and the cloud.

Arsalaan Kashif

Sure, thanks Ritesh. Now, let's really move on to-- from you describing a lot of the components of distributed computing over here, which was definitely useful, I'd like to move on to what was the protagonist of our conversation is, and that is all about how do you address some of these challenges. And I'd like to double-click on the idea of distributed computing, and see how it might be felt, because it seems to me like a lot of the ways in which Computer Vision is being used even today is at a relatively small scale. So for it to really break through into the mainstream and be adopted as widely as, let's say, cloud computing has been today, what are some of the things that you feel need to fall in place and is distributed computing really the secret sauce that can turn out to be a game changer for Computer Vision to really succeed at scale?

Ritesh Gupta

Sure, Arsalaan, and you touched upon it now. Distributed computing, when we talk about distributed computing, there are multiple elements to it. There are multiple elements to it. There is a device, there is an Edge, and then there is a cloud, and you typically distribute your load onto these different layers, if we can call it.

So now, let's look, go to the basics again. When we talk about Edge, what is an Edge? Now, Edge could be a device itself. It could be a device itself which is equipped with the camera and which has got the compute power, number one. Or it could be an Edge machine. Some of the devices are capable of capturing the video streams, but they may not be able to process it. Now, in that scenario, what do we do? We typically connect all these devices to a central, locally available machine. Now, these machines are called the Edge system. Now, with the improvement on the Edge side of it, when we talk about these Edge, Jetson Nano, Intel Movidius Stick, there are multiple libraries that are available that could run on these Edge devices and help in faster inferences. One of the example is Intel OpenVINO. Now, as we said these Edge, all the inference is to take care of the latency part, to take care of the real-time processing and real-time decisions, all the inferences need to be done onto the Edge side of it. The cloud side needs to be complementary to the Edge. That's the basic definition in the whole ecosystem between Edge and cloud.

Now, Edge needs to capture these data, these video streams, needs to break this video streams into multiple images. As we said, Computer Vision is actually image analytics, if you look at it, and we just talked about it, and that's where a lot of confusion comes in. That Computer Vision is image analytics, but it is not image analytics, but then video streams are broken down-- or what is a video stream? Video stream is nothing but the frames, the frames and the images, and they are-- when we play a video, these are actually the images of the frame played with a certain speed. So we say 30 frames per second. In a second, there are 30 frames that play in a video. Now, in the inferences part onto the Edge side, what we need to do is we need to gather these videos, break it down into the images, and based on these images, the analysis needs to be done onto the Edge.

Now, when we do the analysis, the inferences, the decision could be taken onto the Edge side of it, and based on those decisions, it could actually make the devices, who are capable of executing certain commands, can be done. For example, those autonomous cars that we were talking about. Any analysis that has been done, based on the road signs, the decision is taken onto the car and the action is taken onto the car itself. Now, that's where Edge comes into play. Now, nothing is going onto the cloud. What goes onto the cloud is only the incident that has happened. It all depends upon the application and the use cases that we are trying to build. There are use

cases where we actually need to send all the video, entire frames onto the cloud purely from a storage perspective, or there are use cases where these images, only where there are incidents, get reported onto the cloud for further analysis. Now, there is no latency, all decisions are happening onto the cloud. There is no network that comes into play and hence, the chances of data privacy is minimal.

Now, when we talk about this data privacy, there are certain ways this data privacy can be handled between Edge and cloud. One of them is, when the data goes onto the cloud for the storage part of it, it could be encrypted, it could be masked. Similarly, only the incident data can also be masked or encrypted to be stored onto the cloud side of it. Now, this is the Edge piece. Now, let's come onto the cloud. What is cloud responsible for?

Primarily, cloud is responsible for storage, and basically the retraining and model building part of it. We need multiple data sets. So basically, to start off a Computer Vision application, to build on it, we need models. These models are actually trained onto the cloud, we should train it onto the cloud, and these models can be shipped onto the Edge devices for the inferences. Now, as we said, one of the challenges is these models doesn't get retrained with the real-time images that are getting captured. Onto the cloud when these images are sent for the storage, some of these images can be used to retrain the model. Retrain it, create the model, and then ship the model onto the Edge. The typical AI engineering that we call. When we talk about this AI engineering, it is about building models, validating the models, and then shipping it using AI engineering.

Now, that's one storage, model building, and then the other thing onto the cloud is basically the analytics. As I said, if there are certain incidents that have happened onto the-- which has been detected onto the Edge, when these images come onto the cloud, these images can be further analyzed, historic data could be seen, and certain trends could be inferred from that data. For example, if I have to take an example, let's say in a retail store-- I'm just taking a pretty example that comes in my mind, is let's say in a retail store, there is shoplifting happening. Now, when the shoplifting happens, the immediate action can be taken onto the Edge device, which is locally present within the retail store. When this data goes for the further analysis onto the cloud, the cloud can bring up the data analytics part, can bring out the trend-- not a trend. Trend is not the right word, but basically a pattern, if that is emerging out of these shoplifting cases. So that's a simple example.

Now, cloud could be AWS, could be Azure, could be GCP. Now, this talks about certain elements, the important elements from the cloud side of it. Now, how do you build applications, the API gateway definitely comes onto the rescue. We expose these inferences, we expose these data using an API gateway, and the mobile application/web application can be built on top of it. So that gives an overall picture of how we need to distribute deep processing, distribute the storage, and build a scalable distributed computing for Computer Vision. I hope, Arsalaan, that answers your question.

Arsalaan Kashif

It absolutely does, Ritesh. Thanks for the detailed breakdown of the different components of distributed computing and how it can really make a difference.

Now, changing gears just a little bit, I'd like to focus more on the business side for a little while and understand the patterns related to the adoption of Computer Vision, because I know that it has broken through into a few industries in particular, and really understand out there where the market opportunity is, because according to most estimates, it’s expected to be around an $11 billion to $15 billion market and it is being touted, as often the case is with a lot of new technologies, even the ones that have been around for decades, as the next big thing.

So Srikant, let me bring you in here at this point and pick your brain a little bit about this. There are obviously some industries that are likely to benefit more from tapping into Computer Vision than others. So can you walk us through some of the most popular applications of Computer Vision across different industries?

Srikant Sowmyanarayanan

Sure, Arsalaan. So Computer Vision can be applicable across a wide variety of industries, so I'll just touch upon a few of the possible use cases over here. But the applications are not just limited to this. It spans across even other domains, including in areas like automated proctoring in the education space, or it could be telecom tower inspections, etc.

So, in most of the cases, Computer Vision today is being used primarily from a surveillance, security, and inspection perspective. So if we take, for example, the first use case on plant inspection, now, typically in a factory or in complex manufacturing environments like oil and gas rigs, these are pretty large spaces with complex and expensive equipment, doing each of their processes. Now, there are also areas like boiler plants or chimney stacks, etc, which are very hard-to-reach areas within the plant, and very critical to the functioning from a safety perspective, as well as efficiency perspective for these plants.

Now, we could potentially use a drone, which can actually fly around, including inside these equipment structures to carry out inspection by capturing the video streams, and then these video streams can be then analyzed through Edge gateways installed across the factory floor, so that you can take timely actions in terms of predictive maintenance or shutting down certain parts of the plant to avoid critical damage, so on and so forth. You could also actually use this in terms of doing testing and regular maintenance activities as well. For example, drones can be guided using the images which are coming, so a human can actually sit and guide it to do the cleaning of the chimney stacks, for example. So these are some of the examples which could be applicable on the factory floors.

Now, if we go more onto the building side of things, so there are large commercial buildings, who typically already have CCTVs installed on their premises, but these CCTVs today primarily are being monitored by human beings sitting in a centralized command room kind of thing, or a control room, and they really do not leverage anything on the Computer Vision side. Now, you could actually, potentially, automate the inside generation from these multiple CCTV streams to detect the unauthorized users within the building, or you could even detect spaces which are unsafe, and so on and so forth, by combining various building-related data with the visual data, which is coming in through the CCTVs. And that basically helps in doing the better worker safety for these buildings.

Now, Arsalaan, if we move on to the next, from a parking lot perspective, and again, one of the key aspects of a smart city implementation is basically being able to provide smart parking solutions to the residents of the city. And to do a smart parking solution, basically, you need to be able to automate the entire process from the person booking the slot to coming in, the boom barrier, recognizing the car and the person, and then automatically opening up, and then guiding that person to the parking spot. Now, the Computer Vision is used in multiple places. So for example, doing license plate recognition, so that the boom barrier can be operated. When the car is parked in the slot, for example, doing the periodic inspection to detect if there is damage that has occurred while it is sitting in the parking lot, or if there is somebody trying to steal the car, so on and so forth, so that it's a complete smart parking lot with improved safety which can be implemented.

Moving onto the last case study, the last use case that I wanted to talk about over here, is more from a retail and shopping perspective, where the CCTVs can be used to actually help both the customers, basically the people who have come to shop within the store, as well as the retail store associates themselves. Now, for example, if say a shopper is stuck in some place, right? The traditional solution relies on a retail associate walking around and seeing somebody stuck and then approaching that person. But for example, if we use Computer Vision over here, we could actually do this analysis on the video streams which are being captured by the CCTVs and automatically tell a retail associate that, A, there's a shopper in aisle 9 who has been standing there for five minutes, maybe that shopper needs assistance, so you direct the retail associate to places you detect as a possible lead for assistance.

Secondly, you could use it for automatic inventory, stocking and automatic shelf replenishment. So you basically monitor the inventory on all the shelves, and whenever it goes below certain threshold levels based on the video analytics, you can actually then redirect the store associates to stock more of that product. And further on this information can be actually parsed onto the back-end so that the ordering systems also can accordingly read what orders to be placed. So it is not just about the Computer Vision, but then it is married together with your back-end processes. It creates much more powerful business value for the users.

Arsalaan Kashif

Sure, a lot of relevant use cases over there, and especially in the kind of industries that I think a lot of us could relate to on a day-to-day basis, and it's not a secret, of course, that we at Happiest Minds have been very active in this space, and we've been doing a lot of work related to Computer Vision with our clients. I know that we've built a drone-based surveillance platform. We've been doing a fair bit of work on smart robots that is applicable to the retail industry. So Ritesh, if I can ask you to perhaps talk through some of the work that we've done for our clients and bring out some of the interesting elements from it, I think that would be tremendously beneficial to our audience.

Ritesh, can you hear us?

Ritesh Gupta

I'm sorry. So yes, sure, Arsalaan, I'll touch upon the two case studies that we have built in Happiest Minds. The first one is a drone-based real-time surveillance and security system. This was done for one of the South Korean multinational conglomerates. They are into retail, they are into construction, they're everywhere. Now, what we did for them was a surveillance system, drone-based surveillance system, which typically tracks a human being real-time, it detects a human being and tracks the human being throughout its journey or wherever it's working.

Now, typical use case, that typical use case that we were looking at, or the customer, Korean company, was looking for, is basically building the surveillance system into their construction site, into the playgrounds that they have, in the major road intersections. So the whole idea was to fly these drones in an area, and then keep track of the human beings out there.

Now, multiple things that we implemented into this is, since we had only drone or Computer Vision, or only the video part of it, to track a human being, we needed a GPS location of that human being as well, but that was not possible. That's typically not possible when I'm flying a drone independent of that person, or person not knowing it, for example. Now, what we did was typically use a drone monochrome camera, use the drone's GPS location, and along with the altitude, these three, four data points, telemetry data points along with the Computer Vision, we actually calculated the GPS location of a person in a park, in a construction area, or wherever it is. The whole idea of the customer to use this from the trauma safety perspective, also, the workforce safety, when they were trying to do it for the construction site. Now, if there is any accident that happens, a drone would be able to detect it and can create unnecessary alarms.

Now, coming back to the technology part of it, the video that-- we equipped a drone with just a Wi-Fi, cellular-based connection. Drone used to capture the video data over the cellular data, it transmits the video onto the cloud. Over here, we didn't use the Edge part of it so much, but then we transferred the data onto the cloud. As I said, the data frames that reached onto the cloud was close to 20 to 25 frames per second, and we needed lot more. Now, we are still working on it, and maybe we will bring up the Edge-based compute onto the drone part. That's one we will do.

So we implemented multiple object tracking onto the drone's 3D localization. We could achieve close to one meter of MAP, mean average precision. So calculating a GPS location based on a Computer Vision, the telemetry data of the drone, there was a little error that we saw, but we could achieve typically one meter of mean average precision. Now, multiple technologies used over their drones was based on-- the drone was equipped with a single RGB camera. They used FlytOS. They used RetinaNet50, Faster RCNN algorithm, and we did the localization based on the altitude meter, the barometer that drone was equipped with. So by this, what we implemented, the customer was able to use it for their playgrounds, and also into the road junction. Now, from the construction side of it, there are multiple use cases that we are still building on top of it. So that's one of the use cases or case study or work that we have done in Happiest Minds.

The next one is on the retail side of it. So on the retail side, we touched upon multiple times during this session, and this is one of the case studies that we have where we build a semi-autonomous robot. Now, you can see this robot onto the picture. This was a robot that we built, and it was semi-autonomous secured with Mecanum wheels, LIDAR sensors. In this particular case study, we actually did an

Edge processing, and we equipped this, or we had the Intel Movidius Stick onto our drone, which used to do the inferences. So the camera was installed. It used to walk around into a retail store. It would take a picture of the aisle, it will take the picture of the rack, it would count the products on the shelf, and with a certain threshold, if the number of products are reducing, it would take the decision onto the Edge, and once the decision is taken, that notification was sent onto the cloud so that it reaches the store manager, and store manager used to refill or replenish that stack with a product. We are yet to integrate this with the ERP, as Srikant talked about, where it can place an order automatically. So that is still in the works at this point in time.

Arsalaan Kashif

Sure, thanks Ritesh. A couple of very interesting case studies, for sure. Now, as we look to wrap up this and bring this webinar to a close, what I'd really like to do is leave the people who might be listening to this with something that is very tangible and prescriptive, that will be beneficial in terms of ensuring success if they are to dabble with Computer Vision. So let me pose this last question to both of you. When an entity is starting off with its Computer Vision journey, what do you think are some of the best practices or recommendations that they would need to keep in mind in order to truly maximize the value that Computer Vision has to offer?

Ritesh Gupta

Sure, I'll start off and... So, recommendation we touched upon again. Now, the recommendation is, as we said, the decision in Computer Vision in most of the use cases is no more-- is real-time, it is no more just for analytics and no decision to be taken. Now, when it is real-time, as we talked about during our component, when we were explaining the multiple components for the Computer Vision, the inferences should always be done onto the Edge side of it. That should be the first approach, while model building can continue to be on the cloud, but then inferences should happen onto the Edge.

Now, when we talk about Edge, these applications, Edge is distributed. When we say distributed, I'm comparing it with the cloud. The cloud is centralized. When you go onto the Edge side, it is actually distributed. Multiple locations, these Edge devices are there. The other recommendation is all the applications that one builds, all the inference logic needs to be containerized so that it becomes easy to manage, deploy over the air. All those things is possible.

Now, the other thing is, we talked about the inferences on the Edge. Cloud use it as a complementary in terms of doing data analytics on the cloud, the model building onto the cloud. But inference, again, stays onto the Edge side of it. AI engineering plays a very important role, and one should look at the whole process of AI engineering from day one. It shouldn't be an afterthought. It should be thought through when somebody starts their journey onto the Computer Vision. How the model will be built, how the model will be validated, how it will be shipped onto the device, how the retraining will happen. So the whole AI engineering needs to be thought through from the beginning itself.

The other one is on the security side, when we talk about-- we touched upon it, right? The data privacy is very important, a lot of discussions still happening. In multiple countries, no decision has been taken, but then one needs to be sure that the data that is stored needs to be encrypted, needs to be masked. The PII data needs to be taken care of. Security becomes very important.

Now, the other thing is, when we talk about Computer Vision, the algorithms, there are multiple algorithms available-- Faster RCNN, CNN, multiples of them. It is always good to take a call on using an algorithm which suits the use case. Now, this at a high level is the recommendation, which we see, one needs to take care of when they start their journey.

Srikant, if you want to add anything.

Srikant Sowmyanarayanan

Sure, I'll just add maybe a couple of points over here, Ritesh. So on the cloud part, so as we use cloud as a complementary, it is also important to take care of ensuring that Computer Vision does not operate in silo to the business. So doing the right-- linking the insights to your business processes, especially on the cloud side, becomes important to draw out the right value. Secondly, on security,

again, there are some emerging technologies, like for example, the generative AI piece, which can be actually used to generate training data for the models. So that at least solves some amount of problem of handling PII data while handling training data sets, but on the actual stream, we still need to ensure that the security and the data protection measures are put in place.

Arsalaan Kashif

Sure, thank you for that, Ritesh and Srikant, both of you for really outlining recommendations and suggestions, which are very actionable, and are likely to be useful when it comes to leveraging Computer Vision to create meaningful business impact, and hopefully also achieving success at scale.

So to all of those listening in, we hope that you found this conversation to be useful and informative, and if you'd like to get in touch with us, then our contact details are on this particular slide. Do feel free to drop in a line, and we will try to get back to you as soon as we can. But for right now, thank you very much, and we would now turn it over to Lilian for any of the questions that might have come our way through the chat.

Lilian Veras

Excellent, thank you all for sharing such great information with us. We do have a few questions that have come in while you were presenting. So let's get started on those. The first question asked is, can you build a base framework for Computer Vision that can accelerate the development of business apps on top of it?

Ritesh Gupta

Yes, so I'll take a shot at it, Lilian. That's definitely possible. That's definitely possible. Now, if we look at the pipeline, if we look at the whole Computer Vision, we see it as a pipeline. Now, there is an input, there is an ingress of the video, there is a detection algorithm, there is an inference algorithm, and then there is a business use case. Now, the entire pipeline can be built as a framework, and based on the business needs, the inferences, the detection logic, and the business application could be plug and play. So at Happiest Minds, we have built one such framework where, depending upon the use case, we build our detection algorithm, we plug it in, and it's all plug and play, drop it into a folder and the pipeline picks it up, and then there are business logic and the whole pipeline is connected through a messaging system. One which we used was Kafka. But then to answer the question briefly, building a base framework for a Computer Vision application for multiple use cases, that's definitely possible.

Lilian Veras

That's great, thank you. Another question that came in is, can you touch upon some of the pitfalls of using distributed computing with Computer Vision?

Ritesh Gupta

Now, I think we briefly, again, touched upon it. Distributed computing, when we talk about this distributed, these Edge devices, these Edge devices are remotely located onto the premise where these devices are equipped. Now, managing those Edge devices becomes really complicated, and that's the reason one of the recommendations that we talked about is basically doing a containerization of the application. But then management of this Edge is one of the major problems that people see in the distributed environment. However, certain best practices, certain disciplines, if it can be adopted, then these problems or pitfalls of the distributed computing can be taken care of, to a certain extent.

Now, the other thing, which one also needs to think about, it becomes very important, or it becomes very easy on a cloud infrastructure to scale up the system vertically, horizontally and vertically, but on an Edge device, to get more compute, more memory onto the device, it may take a little longer period of time, and hence, the capacity planning onto the Edge becomes important. So yes, the distributed

computing for Computer Vision between Edge and cloud solves most of the problem, but it brings off-- there are certain trade-offs, but then those need to be taken care of.

Lilian Veras

Awesome, thanks for that. Another question here, a member from the audience asking, how do we optimize the use of cloud in the distributed environment and integrate it seamlessly with other business processes?

Srikant Sowmyanarayanan

I'll attempt to answer the question. We actually had referred to that. So in terms of the architecture, where we say use cloud as your complementary resource, while the actual processing takes place on the Edge. So basically, use the cloud for training your models, so that they can be run in a much more highly efficient, modern way, and the models get trained and then deployed on the Edge. Secondly, use cloud to integrate with your business systems, etc, so that the insights can be passed on downstream to other systems so that they can also leverage on the same insights and take action. And third, use cloud as an on-demand system, not like a fixed infrastructure, so that you can optimize the cost of the cloud as the infrastructure when you are running Computer Vision, because otherwise, the infrastructure cost could go quite high. So use it on an on-demand structure.

Lilian Veras

That's great, thank you. We do have time for one last question. A member of the audience says there are very stringent regulations related to data privacy that are being implemented across the world, like GDPR and CCPA to name a few. The use of Computer Vision leads to capturing a lot of personal information that can be sensitive. How does one ensure the highest levels of data security and privacy to avoid misuse or non-compliance?

Srikant Sowmyanarayanan

Sure, so again, I think this is something very interesting over here since we are handling images and video streams. These are typically not stored in your traditional databases. Either these will be in block storages or file systems. So, you actually need to build in protections around your file storage systems as well as block storage systems, and control mechanisms, which is providing users access to this.

So, one way of doing that is basically allowing systems the right roles and privileges to access this data, while you limit the access that humans actually have to this data. Secondly, basically when we talk about, say, GDPR, there are multiple different things like consent, there is also right to forget etc, so you basically need to be able to link this file to a particular user or a set of users and have the audit trails enabled so that as and when these requests come, you can easily go and-- through using the audit trail, you can trace it and delete the data from your storage systems. And then you also follow the best practices in terms of saying implement your right kind of cybersecurity controls so that when applications are accessing this insight, they don't actually have access to the raw data. They can only access the insights. Then you look at putting in a certain amount of regulations to prevent hacking of the servers, etc, so that you follow those specific industry standards, so that you take care of the security and the compliance from a data protection perspective.

Lilian Veras

That is great, thank you. Well, thank you all, once again, for such a great presentation. Thanks to our live audience as well for joining us today. Please do not forget to give our team a rating for the live recording so that we may continuously improve the quality of our webinar.

This concludes our webcast. Thanks, everyone. Bye for now.

Arsalaan Kashif

Thank you, bye.

Ritesh Gupta

Thank you, bye.