EPISODE 54

[INTRODUCTION]

[0:00:10.4] SC: Hello and welcome to another episode of TWIML talk. The podcast where I interview interesting people doing interesting things and machine learning and artificial intelligence. I'm your host Sam Charrington.

The show you're about to hear is part of a series of shows recorded in San Francisco at the Artificial Intelligence Conference which was hosted by our friends at O'Riley and Intel Nervana. In addition to their support for the vent itself, Intel Nervana is also our sponsor for this series of podcasts from the event.

A huge thanks to them for their continued support of this show. Make sure you check out my interview with Naveen Rao, VP and GM of Intel's AI products group and Scott Apeland, director of Intel's developer network which you can find at twimlai.com/talk/51. At the AI conference, Intel Nervana announced a dev cloud.

A cloud hosted hardware and software platform for learning, sandboxing and accelerating the development of AI solutions. The dev cloud will be available to 200,000 developers, researchers, academics and startups via the Intel Nervana AI Academy this month. For more information on the dev cloud or the AI Academy, visit intelnervana.com/devcloud for machine intelligence.

Our first multi-person interview. I speak with Mo Patel, practiced director of AI and deep learning and Laura Froesch, data scientist of Think Big Analytics. Mo and Laura joined me at the AI Conference after their session on training vision models with public transportation data sets.

We talked about this and a bunch of other interesting use cases they worked on involving image analysis and deep learning including an assisted driving system. We also talk through the practical challenges face when working on real machine learning problems like feature detection, data augmentation and training data.

[INTERVIEW]

[0:02:12.3] SC: All right everyone, I am here at the AI conference and I am with a couple of guests this time, I am with Laura Foesch and Mo Patel with Think Big Analytics. In fact, Mo and I, we had an opportunity to meet at a conference a while back and you kind of came up to me and introduced yourself as a listener of the podcast which I think that was maybe the first time that ever happened to me and I was like, excited out of my mind and –

[0:02:39.4] MP: Yeah, I remember you were actually talking to the people at Chainer ,the deep learning framework and I was like, "I recognize that voice" and then it's funny you know? It's a very recognizable voice, yeah.

Then we had an interesting conversation about the industrial AI stuff that I was working on at the time and some of the work that you were doing there. Then when we saw that you and Laura were doing a presentation here at this conference in San Francisco, I thought, "We got to get you on the show," welcome.

[0:03:08.8] LF: Thank you.

[0:03:10.0] SC: This is actually the first time I'm doing an interview with two guests so it would be interesting. To just how the kind of traffic management works. Why don't we start by having Laura, you introduce yourself?

[0:03:23.2] LF: Okay, currently, I'm a data scientist working with Think Big Analytics which you just mentioned. I work on all types of projects. Whenever a customer or a company has a lot of data that they went to gain insight from to solve some use case, I can help them out, it doesn't have to be deep learning, any sort of method that tries to reveal relevant patterns using some method that make sense.

Given the use case and the data at hand and will go with that. Before joining Think Big, I spent half a year in a research group where they investigated something called nonspecific effects of vaccines which is basically vaccines, turn out to affect the immune system in the general way, not just protecting against the targeted disease. Very interesting research.

Prior to that, I had a PHD at the technical University of Denmark using various machine learning techniques to analyze brain activity data.

[0:04:22.3] SC: Wow.

[0:04:23.9] LF: That's sort of my background.

[0:04:25.4] SC: Okay, Mo?

[0:04:27.0] MP: Yeah, I am currently the practice director for AI with Think Big Analytics, mostly looking at America's customers and part of that is probably not as working and projects as much but doing more of the proof of concept type work.

Taking some of the most advanced things that are going out there and see if we can apply them to our client's problems. Part of that you know, there is a hands on portion of it but then there is also the dreaded, like power pointing of things, version of like that like the highly technical stuff into things that people understand which is a kind of a fascinating part of it because I really love that.

Trying to lower the barriers to – because you know, there's a lot of hype around, I try to lower the barriers so that people can understand that this is not terminator, you know, it's actually just math, right?

That's kind of my day to day, I really like doing that and my background is – I come from, you know, if you look at like data science AI, machine learning type things, I come from more of the computer science side compared to like people come from statistics or maybe from some sciences, the hard sciences or the years, the software engineering and to transitioning into more math, the type of software engineering and then do analytics and yeah.

[0:05:38.4] SC: Nice. The two of you did a talk yesterday, it was actually a tutorial?

[0:05:45.5] MP: Yup.

[0:05:46.6] SC: What was the tutorial about?

[0:05:48.4] LF: The tutorial yesterday was an image analysis using deep learning methods in brief. We had both the general introduction making sure that everyone was on the same level, everyone agreed on the image, what are pixels, what sorts of values are we dealing with and so on and so forth.

Also, going into what sorts of problems can we talk about in image analysis and then we went into more detail and one particular topic, object detection and soon bag out a bit in the end. That's how I sort of see the whole framing of the talk, I don't know if you have anything to –

[0:06:30.9] MP: Yeah, absolutely. I think just to kind of making sure that people are aware of the computer vision basics and then diving into something that is fairly cutting edge, object detection and a lot of applications out there and not only the theoretical part but also you know, in a notebook style kind of layout saying, "Hey, this is how we can actually do it yourself as well," right?

I felt that was very compelling and then towards the end, talked about some of the challenges around training models, you know, it's like, we make it sound so easy but to do it for real data, either there are many challenges and so we talked about that and we can go to that detail if you want.

[0:07:09.1] SC: Yeah, what are some of the challenges?

[0:07:11.6] MP: Absolutely, you know? For example, there was a paper that came out or the weekend, how this team trained image net data set in 24 minutes, right?

[0:07:23.4] SC: A little bit of a controversy.

[0:07:24.2] MP: Yeah, of course yeah. Because you know, it's like well they used – that headline was Alex net not res net which is like more of the current state of the art and what kind of hardware you use and all sorts of different things but that is exactly the type of thing, right?

What are all the different parameters like bat sizes and different processors and multi GPU, multi server. Because most of the things you see the examples of tutorials, it's like just around this code, right? If you're in production, when you have like million images and you need to make sure that this will train days and not weeks, you know, you may have to scale it and how do you do the scaling and those are all the challenges that's kind of like more engineering side challenges but then there's also challenges around annotation, right?

Well, this is supervised learning, you have to annotate the data and that could mean anything from, it's like, simple classification, kind of easy so to speak, right? There's a picture and there's a label, right?

Now you, for object detection, there could be bounding boxes so you draw squares around objects like around balls and you know, like umbrella and things like that to make sure and then label it this is what the object is.

Then even something more advanced which is kind of drawing the polygons around the object itself which is a segmentation, kind of like the holy grail of getting towards being able to do object detection.

Many challenges and of course, you can try to do it internally or externally, we actually – for our project, we actually built a segmentation tool that allow our people to go ahead and draw boxes around cars and pedestrians and things like that.

As much as we talk about the deep learning part, there's all sorts of many data engineering, data prep, data cleaning, things that have been around and also nutritional data science for a while, right? very much of a challenge.

[0:09:14.9] LF: And to add to that a bit with the annotation tool that we build internally, we have that tool and we were using it internally but our team was just not large enough to annotate enough images quickly enough so we had to both use that and go to an external company to help get help from them to annotate all the images.

That was a lot of back and forth with them just defining the requirements. “What sorts of things do we want labeled? How do we want them labeled? What’s the smallest size of object that we require labels on?

What do we do if it’s partially obscured? If there’s a car blocking another car?" Those sorts of things because you may have difficulties if you have a really small object and you label it then during the training phase, you’ll be punished, the metal will be punished if you don’t detect that object but it may actually not be very interesting to detect that object at test time.

Because small objects are far away. You may want to focus on updates that are closer. One way to handle that would be to put an extra label and small update saying difficult.

That way you might handle such objects differently from normally labeled objects. Punishing the model if it doesn’t detect them but also not punishing the metal if it does detect them.

[0:10:30.2] SC: Yeah, I’ve had some interesting conversations with folks that specialize in labeling data for folks and it really opened my eyes to just this process that you’re describing like – you think of “Hey, just label my data” but if you're talking about images, all different kinds of ways that you can do it.

They have direct impact on the types of, you know, not just the types of models that you’re creating in a performance but the cost of the labeling process.

[0:10:56.4] LF: Yeah, so this company that we worked with had this quite elaborate pricing scheme, I never really looked at the details of it but if you increased the number of classes, you would sort of get an extra cost to the first classes as well so you really had to consider like maybe can we sort of have this external company part of the labeling and then do some further post processing in our own tool.

[0:11:17.9] SC: You needed a machine learning model to optimize the process for the vendor tool?

[0:11:22.6] LF: Yeah, exactly.

[0:11:24.1] MP: you know, there is some interesting projects around that, either there's that snorkel project from the Down Lab at Stanford they are trying to do some stuff around kind of, around the training data and building more training data which just reminded me another big one is the data augmentation which is another thing that you know, if your data doesn't have, example I gave is that, having the raw data for when it's foggy versus when it's not, right?

What if we never captured the fog base which is harder to do in San Francisco but you know, those are all the type of things that – luckily, you know, at least what's great is that not much of this knowledge has been encoded now where they're just in curious.

They're just a function for data augmentation. Sure, maybe you could add your own but there is that state of the art, like might as well just use that when you're in your training process but these are all the things and it's like, when you think about it simply, yeah, just take the images, they don't have to deep learning model and out comes your train model after you do all the deep learning things like generalization and lost optimization, all those thing.

But then, there is all these other things that you have kind of worry about.

[0:12:32.1] LF: Regarding data augmentation, even though people have made tools that you can just use your – of course have to think about all of the data augmentation steps to them, it makes sense, for traffic scenes, do you really want to flip your images horizontally to teach the model that an upside-down car is also a car.

You do but it depends on your use case, right?

[0:12:52.7] SC: Right, that will depend on your use case.

[0:12:54.4] MP: Yeah, I mentioned the action movies, you know? Get those upside-down cars, yeah.

[0:13:01.5] SC: These are just some of the issues that you run into in the training phase of building and deploying a deep learning model and then there is the whole, how do you actually get this out into the wild to do imprints like are there best practices, tools of the trade or you know, tricks that you've come across for that?

[0:13:20.6] MP: Yeah, I mean, actually, as much knowledge there is out there about the training aspects of it, there's actually not as much and something that – even traditional machine learning in production, I think we heard this thing about data science and there's not a – there's a lot of talk about it but end production is still not – you could do a survey and you could probably find that only maybe of all the people who actually say they do data science, only 25 to 30% maybe actually putting into production and then once you get to deep learning that, that number could start even going lower.

I think that's just because as you were saying that there's a whole different set of challenges when you are trying to put models into production, right? Number one, where are you going to put it? Are you going to put it on some beefy servers in a data center?

Then maybe you have lesser problems because you could take those big gigabyte sized models and democrat containers and kind of the traditionally of apps are served. But then you start talking about, "well, we're going to do mobile," that becomes another challenge, how do you compress the models, maybe quantization, lowering the floating point and things like that. You know, that's another –

Somebody actually asked this question, they're like, "Can I take this in and put it into a cart to do this," we're like, "There is lot more that would go into it before you'd be able to do that," the training is definitely challenging but being able to serve the models at scale brings in kind of all your traditional dev ops and data ops kind of bring it all in.

Model management, version control of the models and data lineage, addressability. Everything that we've been discussing for any other data science type things, those are all – they are back on the table, right? How complex those can be.

[0:15:01.7] SC: Are there emerging or accepted tools or open source projects for doing that kind of thing? I know that you know, some of the folks, some of the vendors that focus on machine learning platforms, they've got some of that stuff built in or integrated into their toolset but it's all like you know, within that tool set, are there, you know, is there a kind of an open source model management framework for example that's kind of emerging as a standard or does it have to be kind of custom cut for an individual use case?

[0:15:32.6] MP: At least I have come across, I know we built one internally, yeah, from just a lot of our projects, right? Because we saw this need of model management and you know, it's internally called think deep, you know, for managing models, it's maybe it will become open source, you know? You know how these things go.

At Think Big, we have tried to make things open source, for example, the Kylo framework for data links, this could be another path on that roadmap but you know, there's a lot of polishing that needs to be done before it gets.

[0:16:02.4] SC: Reach out to you if they want to get their hands on this open source management?

[0:16:06.5] MP: Possibly, I'm not the offer on it so I can definitely put in touch but I have not come across, they're definitely some projects once again, the formerly amp lab, the UCB rise now, right? And then, a Stanford don.

They have, I can't remember the names of the projects off the top of my head but I've seen that in their program agenda for being able to serve large scale machine learning models and you could consider it deep learning into that.

I'm not due to a commercial plug but that's actually one of the things that we are also focused on. Because when we look at things, it's like yeah, training is there but we work with traditional customers, enterprise customers who want to put all that stuff into production because all the investment they make on AI or deep learning, data science is useless unless they actually put the models into production, right?

Building the tooling around that, monitoring the models and all of those things are inter ability is a huge part, all of those being able to kind of have a one stop shop for that is very attractive, yeah. There is commercial aspects to it.

[0:17:13.6] SC: Interesting. Before we got started, we mentioned a few use cases that you will be able to kind of talk to us about and walk us through and the first one is one that you worked on Laura as a traffic project that you worked on with an automotive parts manufacturer and it sounds like this was in some ways an inspiration for the session, the tutorial that you did here? Can you tell us about that?

[0:17:37.5] LF: This project was for an assisted driving system, not self-driving at all, we didn't want to go there so help cars help each other essentially. The idea was that you have a car that's connected to the system and it's driving along like doing whatever it's doing, going where it's going and it has a camera recording, whatever it passes.

If it happens to pass a stopped car that's on the road, it might be a good thing to be able to detect that, to tell other cars, "Hey, look out, if you're in this lane, you might want to change lanes already now," you don't sort of get a surprise when you get to that car.

You could even have that indicate there's an accident or congestion. Stop cars on the road, it's definitely something that you might want to tell other cars about. The project was about trying to detect this.

There are a lot of steps to this. First, a video, of course a lot of images that come in one at a time so the first step was trying to look at, "okay, how well can we detect the objects and trying to compare the various object detection methods that were out there, see how fast are they, how accurate are they," all of these things and also try to get an idea of how might we be able to improve them.

Then subsequently, "you have to be able to tell us this part of the picture, part of the road or is it not part of the road?" Because if it's not part of the road, if it's parking lot, you don't care so much whether there's stopped cars in there. That's another thing.

Thirdly, you'll want to be able to tell, "is the car moving?" One thing is detecting the cars in each image, another thing is seeing – this is the same car that I just saw in the previous image and then estimating the distance, the difference in distance from when you saw it last, trying to estimate its speed.

This comes into object tracking which is think too from the research that we were able to do to be still quite immature compared to object detection. For object detection, there are already a lot of methods, a lot of research out there for object tracking, it seemed to be less solved.

We spend some time looking into that. In the end, we had the demo that was able to detect most cars and sort of give an estimate of is it moving how fast is it moving? That was basically where we stopped.

[0:20:14.2] SC: Okay, were you exclusively trying to detect cars or other obstacles in the road?

[0:20:20.5] LF: We did want to detect the object in the start of the project as well, turns out that there were some difficulties with that because in traffic data sets, you tend to have a lot of cars.

You don't have as many people, you don't have as many bikes, you don't have as many buses so your training data has a high imbalance and that just makes it more difficult to learn those other classes that are not cars. In the end, we ended up focusing more in just the car part of it and not so much on detecting all the other classes that we were initially interested in.

[0:20:54.9] SC: Okay. Interesting. You mentioned the object detection methods that you came across, can you kind of summarize this data of object detection and what are the main methods and what you found as you compared them one to the other?

[0:21:12.6] LF: The methods that we looked into most were the ones that were quite recent at the time of the project which was right in the beginning of this year. At that time, there were a lot of, what's referred to single shot methods so you have two versions of one called, you only look once, YOLO, there's one called Single Shot Multi Bug Detector and what these methods have in common is that they just look at the image once, you'll take the image that you're trying to analyze and pass it through the middle ones to extract features.

Then based on those features, you'll predict the coordinates of the bounding bugs as well as the class of the object within the bounding bugs. You'll do that for a bunch of predefined bugs in the image. Before like seeing the image, you'll already have defined a large number of what's called Prior Boxes or Default Box System, they have various names.

For one particular method, it's around 7,300 boxes. For each of these boxes, you give a prediction of the class that it might contain and a correction to the predefined coordinates to match the object better and that's sort of the general theme for these single shot methods.

[0:22:29.1] SC: Okay, is that they generally they predefine some large set of boxes and then detect objects relative to those boxes?

[0:22:37.8] MP: Yeah, exactly.

[0:22:38.5] SC: Okay, interesting. Then, on the object tracking side, what was the method that you found and how did you – what was your experience with that, how did it perform?

[0:22:49.7] LF: We ended up using a heuristic approach, so basically looking at the detected objects and a number of frames I think we ended up with looking back at the past 20 frames. Basically trying to match the current card that we're looking at to the car in the previous images that matches the color the best and has the closest Euclidian distance and those sorts of heuristics. That turned out to work pretty well.

[0:23:17.4] SC: Okay. You mentioned kind of doing a scan of the literature, was that one of the methods that you found in the research or will you not able to get any of that stuff to work well?

[0:23:29.1] LF: Yeah, exactly. The methods that we found and tried out, we spent quite a while trying to get them to work and it didn't really work out so in the end we just thought, "Okay, let's try heuristically, see how it works" and it ended up working pretty well.

So then, in the interest of time, we went with that and of course we're hoping to get some of them more advanced methods to work in the future but the heuristic approach turned out to really give quite good results.

[0:23:58.1] SC: Which is also an important lesson for folks that actually have problems to solve.

[0:24:01.9] MP: Yeah and you know I was observing that project mostly and of course as you were just saying, for an outsider I was like, “Oh this is clearly, you need to use, they are using color wishful networks but then like use recurrent on top of that because you are trying to track something and then maybe you do the predictions for future frames and things of that sort,” right? And that was the common and of course that’s what I search for right then. There is somebody who is working on a recurrent YOLO.

[0:24:34.8] LF: We did try that actually. That was one of the things we tried and –

[0:24:37.4] MP: Yeah, it didn’t work and that’s what we mean that maybe this is something that will as more people will experiment and it will start evolving because object tracking is another great area for computer vision.

[0:24:49.9] SC: That’s a big challenge for folks that are implementing this stuff like in real use cases is that you know, I’ve heard it described in a repeated as kind of overfeeding on a dataset right? Like, “Hey this works great for imaging that” right? But you know for another dataset it doesn’t. It’s hard to reproduce the results.

[0:25:09.1] LF: And that’s actually another issue that we ran into. So we took this pre-print model that someone had under GitHub Repo just to see how it worked and it didn’t detect anything. So this was then taking model and we wanted to detect routes and it just didn’t detect anything and of course in our images, we had a lot of routes because this was a traffic dataset. So we were like, “Why is this?” and one of the team members wrote the often.

He was like, “Oh well that’s because it didn’t do data augmentation because I wanted to improve the performance as much as possible to the dataset that I’m submitting this entry too” so this was a competition and he was just optimizing to this lighting conditions and so on that the dataset had. So when we retrained the model with data augmentation it was actually able to detect a lot more road even when we didn’t add any data, we just used data augmentation.

[0:26:02.0] SC: And what is data augmentation doing in this process?

[0:26:04.5] LF: So I think some of the perimeters that we churned in this case was suggesting the brightness of the training images. We had the original training images that this guy had trained on but then we added the same images but with different levels of brightness and other things to the training set.

[0:26:23.5] MP: So essentially to deal with things like glare, nighttime, twilight and all sorts of other lighting type of situations.

[0:26:32.1] SC: So you do a bunch of different transformations on the image to adjust the brightness by a few plus minus, a few stops or maybe apply some cool Instagram filters?

[0:26:42.6] MP: There you go. So yeah what's great, once again I love this field and everything that people are doing because there is a paper around this topic of data augmentation and best practices and much of it has been platified and care as some of the other computer original libraries. So you don't have to do imagination. It's like, "Let's just apply this best in class even data augmentation and then we'll just see if it works." Of course if it doesn't then you may have to go inside and tweak some things, yeah.

[0:27:13.2] SC: Okay, awesome. Anything else on that project?

[0:27:16.4] LF: Not that I could think of right now.

[0:27:20.4] MP: Yeah, I think another thing just from my observations was that international like everybody has their local perspective right? And we have a US-based, US centric view point on traffic and data driving and I think there's just different set of things in countries and actually one good example of that is traffic signs. Stop sign, the concept to somebody universal but it is different in other countries.

So these are the type of things that you heard about or come for. If you are trying to do more like sign detection or other types of things.

[0:28:00.3] SC: Excellent. So you mentioned a project called Lost and Found that was for a logistics company?

[0:28:07.3] MP: Yeah and that one as much as this kind of object detection type things that one had a little bit of different type of challenge and that is that type of thing that we are trying to find. So imagine you're shipping packages and then somehow the package loses its tracking. The label rips off or box break and so this company finds item and now it's like, "Okay what is this?" because a lot of times it's not really clearly or you as a customer will call and say, "I'm missing a brown pair of shoes with this, this and this".

And somebody is manually going through all the stuff to find it. Yeah, it is a very tedious process and so how do you build a system that could give you based on some description and I know Google image search this is great but we don't have – it's not available as an API even. Google is not giving it to everybody and so everyone calls API's but still not this. The point I am trying to make is that then numbers of items that you could lose is infinite.

It's not like a 10 class, 90 class, 9,000 class. It's infinite so doing an object detection model that will say, "Oh will just detect the label of the item and we will just be able to find it" so the approach was a little bit of a combination of transfer learning and the future matching as oppose to like future engineering, future learning. So one thing that customers do when they ask the question is they actually could provide a photo from the internet or they may actually have the item.

And this all sounds probably somewhat easy for commercial items right? But if you think about industrial items or niche things or collectibles and stuff like that, that's where you start really getting into the problems of finding these items.

[0:29:51.4] SC: Well certainly you are not going to have them just sitting around in your training data?

[0:29:54.9] MP: Yeah. The way to tackle this problem is like an image search problem, right? Previous generation techniques where doing a lot of the hand crafted features, there are algorithms like SIFT, SURF and score and some of these previous generation feature detection algorithms. I believe –

[0:30:14.7] SC: Sift surf?

[0:30:16.5] MP: Yeah, SIFT, SURF, these are all acronyms for feature detection in previous detection in previous generation computer region. As a matter of fact I think and of course I don't know the details but when I look at it, the Amazon app when you open up the visual search, I think they are not using deep learning. They are using the previous generation because of the dots it's showing and based on just my knowledge that's exactly how it works.

So I don't know if maybe somebody from Amazon could verify it or maybe not. They probably don't want to but yes, one of them I forgot which one is actually proprietary so we can use it without getting a license and all sorts of challenges like that and we have deep learning now. So being able to extract using some of the straight state of the art models such as ResNet or Inception and just really thinking about this problem. It's like, "well what are those models doing" right?

You have high level, high dimensional data images and the point of generalization is to come up with low level features that you can generalized so that it will work on not just that data but on a much wider data. So why don't we just try to take our target and the source and essentially extract those low level features and create a feature store right? So the next time you had to search for something, you say, "Okay let me" it's almost fingerprinting matching type.

That is one way to think about it, right? Creating the fingerprint of the search and then we have the entire database of fingerprints and you just say, "Okay find the one that is the closest match" and we don't have to get it a 100% right because there was a human in the loop that somebody will say, "Okay all right I can tell within this fire images that this is the one", right?

[0:31:58.2] SC: Did you define your features a priory or did the training process?

[0:32:03.1] MP: Yeah, we didn't do any training. We just used the pre-train model for example ResNet or Inception. We removed some of the classification layers because we are interchanging the classification right? If we look at the deep learning model, that's what the lower layers are doing. It's eventually after the features are you get to a low level you then do your classification and then we use that. We are not interested in that.

Let's freeze the model at a certain level and then see what the features are coming out of that. Store those and then let's do the same thing for the search candidate and then the object tracking, there is a lot of similarities as far as the cleared in distance type of thing where you are trying to make sure there is the same car from the previous image to this image and then once you extract the features, you try to do a distance and see how closely they match and that's essentially what you are trying to do.

Is that in this case we are doing a cosign distance between both of the features, the feature vectors or kind of the customers features vector and then the entire database of images objects that we have to find the closest one. So that is kind of like in a nutshell that's how that problem will solve.

[0:33:10.2] SC: On that one did you determine where you were going to freeze your network based on experimentation or was it more kind of intuitive, "this is where the classification starts and we'll just going to stop here"?

[0:33:22.4] MP: Yeah, if we ever dissect one of these models you can read all the layers and you can tell which layers is when the feature extraction stops and the classification starts. Typically around densely connecting in Soft Max Classifier, those are the later stages. So once you stop your illusion layers that's when you can stop. We try to experiment until the very end or somewhere in the middle because these models are also the state of the art.

They have 50 layers, many layers because we were not really interested in doing the classification. So at what level of raw features that our search works in a good performance, right? Because again, if you get to really low level then our search results will be all over the place. So we want to capture it at some good level. Not a scientific way, it's pretty much a trial and error to come up with giving us the best results.

[0:34:24.0] LF: I was actually wondering now that you used the features from many layers at one sort, did you pick just one layer and take the features from that?

[0:34:31.1] MP: Yeah, after several layers like 25 or 30 layers we said, “Okay now let’s see what the output of the image is,” the future vector and said, “No this is the one we want to use”.

[0:34:41.9] LF: Yeah, I mean I’m thinking inspired by these object detection methods that tend to use features from different layers to capture both large and small optics and then make but I guess it is not so important if you know that you have one object in the image. You are just trying to find that one object right?

[0:34:57.8] MP: Yeah, typically what I didn’t mention is that this company when they get the item, they take photos of it, multiple angles and everything. Typically that is the only thing that is in there. Once again, all classic problems are still on there like data cleaning and background removal and all sorts of things like that, yeah.

[0:35:19.0] SC: Yeah, interesting and you mentioned one more which is a fraud detection app for a bank?

[0:35:26.3] MP: Yeah and did you want to talk about it? I can also talk about it you know?

[0:35:30.0] LF: If you could start, I wasn’t on the thought. I know about the project but yeah.

[0:35:35.0] MP: Yeah, I just felt like because it’s in Denmark right? And yeah, I think more details can be found, the video of this project. There was a talk about it earlier at a conference in New York in June and it is available online. So those who are interested into this technique. I think what’s really neat about this is that in the previous two cases we talked about computer vision which is the cool factor because everybody is excited about being able to do things with images.

But now, here comes fraud detection which is to the common person it’s like, “Who cares? As long as my credit card works and nobody frauds me” but banks of course are very much concerned with that, right? So in this case of course fraud detection is not new. Some people have been doing this for decades from all sorts of things like human curated rules. They use the fiber card in San Francisco and then you swipe it in Bombay, obviously there is a problem.

So things like that you can always have those handcrafted rules but those can become 20,000 rules and it becomes a real cumbersome process. You start applying some traditional machine learning aspects to it. So you do a lot of feature engineering on the data to say, “Okay these are the features that contribute to fraud and we should flag for that” right? And then we are like, “Well what are the ways you can improve upon this?”

Now we have deep learning, feature learning, are they ways to actually improve on the model and so we said, “Well let’s try a couple of things” so one of the things and this is a classic way of object tracking example. “Well what is the first thing I would do?” And most people for transactional data or sequential data, the go-to technique is using some kind of further kernel model where LSTM or something like that. So you have that recent history and then you should be able to tell, “Oh this is fraudulent”.

It is about to fraudulent, so we tried that and actually the model result if I recall correctly, there is chart in the talk which shows you all the different approaches. I think it was on par with the traditional machine learning. So not really promising but then it’s like a creative way of, “What if we could apply some of the vision based techniques?” so there are certain properties that are necessary for a panel illusion model to work.

In variant statistics and locally correlated values, those are the two things that famously Yan Macoon tweeted out a few months ago which I remember very well. I am not having come from research and those things really made a lot of sense to me. It’s like, “Okay, all right so those are the properties you need” because traditionally the type of data we have even in credit card transactions is some kind of time series. You have a model with them and location and all sorts of things like that.

But not like a picture and this technique is detailed in you of five minutes of that talk there is a couple of slides on it which visually is very intuitive is to come up with a feature map of the transactions using some history. So then the idea is to create like a visual image off the recent transactions and then feed it through the convolutional neural network to see if you could detect fraud. When you think about this and this is something I would ponder others to think about.

Think about this as how we would do something like that and one example I always talk about is that people who sit in operations controls and they are looking at this 50 screens right? Or even traders when they look at all these charts and all these things, they are visually also trying to look for some changes and signals that would allow say, “Oh I need to take some action” and the idea is very similar here, right? If we could convert there’s data behind those charts right?

If we could somehow create the – and just create charts out of the data and then feed it through the con illusional neuro network and then tell it what abnormal looks like right? And then now ask you to flag it so –

[0:39:32.9] SC: It’s super interesting I think those traders will be a lot less effective if they were looking at a spreadsheet.

[0:39:38.5] MP: Yeah we know that right? It’s like imagine if I gave you like a trend chart versus like all the data in the table. How quickly could you make a decision?

[0:39:48.4] SC: Put like that it’s super intuitive that a neural net would, a vision focused neural net would be effective in solving this kinds of problems.

[0:39:57.0] MP: Yeah and you know I love talking about this because it is a great example of applying some of this, what the research there and the progress that we are making in computer vision towards probably a newer thing is in that realm and you can talk about – when you start thinking about this way there are many more things that open up like anomaly detection and anything that you can essentially visualize like if you have a heat map.

So yeah, sure we can generate a heat map of a lot of data, right? I think the big trick over there is you have to be consistent in the way you feed the data because you remember the properties of the combination of neural networks require that – think of it – it is an image. If you start shifting around the bits the image will stop making sense right? And that’s what we are really banking on that there is an image that makes sense to us and we are trying to find those patterns.

So it is very important that you don't use them, other technique the next time around. You keep the feature app same throughout the production cycle not change it around, not change the sequence of the transactions and things like that.

[0:41:03.7] SC: Okay.

[0:41:04.2] MP: Yeah it's detailed much more in that talk if anybody is interested.

[0:41:08.4] SC: We will try to track down the link to that. Awesome, well these are all really, really interesting use cases. Thank you both for taking the time to talk through them. Is there anything else that you want to mention? Either of you.

[0:41:22.8] LF: Nothing comes to mind right now, probably something will in half an hour.

[0:41:26.9] SC: Yeah, of course.

[0:41:28.5] MP: I'll take Sam's rule and I'll ask you what has your impression been in this I guess day and a half so far at the conference?

[0:41:36.2] LF: Well I have enjoyed it personally. I think there are really interesting talks and some that I would have liked to go to but that we're just filled up by the time I got there. There was one on what to do if you don't have a lot of data because there are some ways around that I gather and something that a lot of people are starting to talk about as well as the use of unsupervised learning. I don't remember the details but I ran into this really interesting paper when they started.

So they wanted to train some kind of image analysis optic detection something method, I don't recall the details. The insight that they had was that well, "how do children learn to associate shapes that belong together?" If a child sees a cat the first time, it may not know that the legs and the head and sort of everything goes together into being a cat but then when the child sees the cat move a lot, it realizes, “Okay so this is a cat and these parts belong together so this is one object”.

So this is essentially learning how to detect a whole object and there is some research that shows that this is indeed how children learn how to recognize shapes or people who regain sight that there is actually some research showing that this is how it might work for humans. So based on that insight what these authors did was just show videos to deep learning model and it did learn to recognize objects in this way and then once of course once it learned how to recognize updates then it is easier for it to learn that, “Okay, so this object is a cat. This object is a house”. So yeah I thought that was a really cool application of using unsupervised learning to gain momentum.

[0:43:23.1] SC: Oh wow.

[0:43:23.5] MP: Yeah like a little bit of what we’re down about way of going around the labelling aspect, yeah it is pretty neat and I think it’s great for some of these research moving forward around because we have been talking about big data related machine learning with deep learning but there is a lot of data, class and balances of huge, huge things. There is some places where it is not enough things to detect how do you handle those things, yeah.

So it is very fascinating and I’ll say that for me today it was really fascinating the keynote from Rendering on a very non-traditional. He broke out the wide wards and I would recommend a lot of people to check out that because I really felt that of course he was really introspective and things like that but the things that he highlighted I really felt they really hit home because of course people give great keynotes and I think they are good but I think he really said:

“These are the real things that you need to worry about or look at that are very relevant versus let’s just try to ride, increase the hype car further and further” right? I thought that was very – that is something that I know I would reference several times watching the video because it really brought some of the key points out of there.

[0:44:37.5] SC: Awesome, great. Well thanks both of you.

[0:44:40.3] LF: Thank you for having us come.

[0:44:41.8] MP: Yeah, thank you. Long time listener and finally hearing my voice, you know it's going to be weird. So thank you so much for having us.

[0:44:49.4] SC: Awesome, enjoy the rest of the conference.

[END OF INTERVIEW]

[0:44:55.6] SC: All right everyone, that's our show for today. Thanks so much for listening and of course, for your ongoing feedback and support. For more information on Mo and Laura or any of the other topics covered in this episode, head on over to twimlai.com/talk/54. For the rest of this series, head over to twimlai.com/aisf2017 and please, please, please send us any questions or comments that you may have for us or our guest via Twitter @twimlai or @samcharrington or leave a comment on the show notes page.

There are a ton of great conference coming up through the end of the year. To stay up to date on which events we'll be attending and hopefully to meet us there, check out our new events page at twimlai.com/events. Thanks again for listening and catch you next time.

[END]

