Deep Learning in Depth: IARPA's Functional Map of the World Challenge

featuring Ritwik Gupta and Carson Sestili as Interviewed by Will Hayes

----

Welcome to the SEI Podcast Series, a production of the Carnegie Mellon University Software Engineering Institute. The SEI is a federally funded research and development center sponsored by the United States Department of Defense and housed here on the campus of Carnegie Mellon University. A transcript of today's podcast is posted on the SEI website at sei.cmu.edu/podcasts.

Will Hayes: One of the most interesting recent applications comes from the Intelligence Advanced Research Project Activity's work on Functional Map of the World Challenge. And you gentlemen were both part of that?

Carson: That is correct.

Will: Tell us a little bit about that because that sounds like a neat project.

Carson: We refer to it as IARPA. If you are interested, you can always check out a link in the transcript of the IARPA Functional Map of the World Challenge. This was an image-recognition challenge on steroids. The problem was the United States has a lot of satellites that are fixed above the earth and has an overwhelming amount of satellite imagery of portions of the earth's surface.

What they are interested in doing is finding out what is going on on these plots of land. There are various functions that can be ascribed to buildings or facilities on the surface of the earth, like airport, amusement park, nuclear facility, or hospital. It is really important to be able to tell what is in this chunk of land for lots of intelligence-related reasons. There were several complications to the challenge that made it more interesting than just the standard, *Does this photo have a dog or a cat in it?* But, it was in spirit a very similar approach.

Ritwik: I do want to add that it is not just identifying the functional buildings, it is also identifying a random piece of land. *So, is this a crop field, or is this a flooded road ?* Again there are a variety of applications for this. IARPA obviously would be intelligence based, but there are also large amounts of humanitarian use cases for this stuff.

One distinct thing I can imagine for a really heavy use case of satellite imagery was there was a city, Fort-something in Canada, which was ravaged by wildfires a year or two ago. One of the things that happened was, when humanitarian workers went in to do rescue operations and to provide aid, they did not know where structures stood anymore. They had to use satellite imagery to identify what…like this square on the map now was actually a hospital before, et cetera. There are a lot of use cases for stuff like this.

Will: So it could help them better navigate hazards. It can help them understand where there might not be anything solid to stand on, even though it is not apparent here where there might be more people in need of rescue.

Ritwik: Exactly.

Carson: There is not a barn here anymore because it burned down, but there used to be one and, we need to know that.

Ritwik: We can tell you that because we know what the function of the land is. Therefore the Functional Map of the World.

Will: Could you talk a little bit more about the collaboration you had on the Functional Map of the World project?

Ritwik: I can start off on that. Our goal for entering the challenge was kind of twofold. One was to basically work on a fun little problem that had real use cases for the United States government. The other one was to basically try out very new methods ourselves and find limitations within our own infrastructure, within our own methodology. There is kind of a two-pronged way. It was not necessary to win the challenge.

Again, the task is, you are given 62 different possible functions of the world, plus one false detection category. That means that this was just bad label data. That task was that you are given about five terabytes of data of which about 65 percent is training data. Training data does not contain false detections. Can you using that data in the best manner possible identify the land use of the satellite imagery? What we did was, we tried a variety of methods to do that: variety of existing methods, deep learning methods, you know, using models such as DenseNets, using models such as SENets, that's squeeze-and-excitation networks.

Carson: So if I can interrupt for a second, a very high-level view of this is that, this is just an image recognition challenge, which is what deep learning kind of came to fruition in proving to the world that it was good at. Image recognition is like the slam-dunk success of deep learning so far. We said, not only we, the people who instantiated the challenge said, *Here's a starter deep learning architecture*. We used a variety of different deep learning methods that have been used for image recognition in the past. Other such challenges were given on the set of labeling all the images that made it on to Google Images. The technology behind this is convolutional neural networks, which is a really good way to get into deep learning if you are interested in learning more.

Ritwik: Again, Carson previously mentioned that this is not standard image recognition tasks, it is not standard image recognition classification tasks. That is because generally when we are talking about categories, like classification, it is like dogs versus cats, or it might be something that's very close, like Porsche versus Lamborghini, or certain sports cars. This one is unique because you have satellite imagery, which for a large part of the earth is very homogeneous, right? One patch of forest in the USA might look exactly like a patch of forest in the Amazon, right? Or a city from like a zoomed-out view might look the same as the next city.

The idea is how can you take these very minute differences, not only in scale, but also in landscape, the buildings on there, etc., and identify different land functions. This makes it very different from just a traditional image classification problem because you have to take in not only the object of interest, which is like let's say a building, but also its entire surroundings.

Will: For example, what a farm looks like in Kansas versus what a farm looks like in New Delhi? They are very different things. Not only are they growing different crops, but the size of the fields, the machines used, the seasonal activity that happens are very different. This challenge needs to be able to accommodate those kinds of sources of variation.

Ritwik: Yes, like crop fields in Kansas have irrigation circles, right? It is a very modern technology, these irrigation circles. Crop fields in New Delhi don't have that. They are usually hand-watered or they have pipes running in the field. On satellite imagery, those would look very different from each other. So, identifying something like that is important, but you do not really get that in just general deep learning tasks.

Will: The feature learning task is what this challenge addresses.

Ritwik: This is the most critical part of this challenge, yes.

Carson: If I can pull that back to our feature representation discussion from earlier, unless you are a world expert in what satellite imagery should look like, or you know a lot about farming techniques or whatever, it could take you years to determine this is what a crop field in Kansas is going to look like, this is what one in New Delhi is going to look like. You have terabytes of data, you don't have enough time in order to make that happen. That was the whole reason that we needed to use this deep learning strategy is because we are not experts on what satellite images look like, but the computer can become one.

Ritwik: It is not to say that we are completely clueless about satellite imagery. We have some idea, which we try to incorporate in the models. The idea is, it can take our cluelessness, or some naive knowledge of it and build on top of it and become really good at identifying all these things.

Will: It is the power of what the algorithms and the techniques are able to do beyond whatever the human brings to it, that we are really trying to test with this challenge.

Ritwik: Correct. I think one of the biggest things is not only what can it do beyond humans, but the computation power and the infrastructure that lies behind the deep learning, which really empowers it. We are really good at doing all these cool abstract tasks, we being humans, like identifying what your grandmother looks like with one look at your grandmother, or reasoning about complex things, like *Oh, that is a bookshelf*', and *Looking at those books, we're probably in a Software Engineering Institute*. What the computers are really good at is doing one thing really well, and fast, really fast, much faster than humans can.

So how we best leverage all that architecture, the infrastructure that has been built around, or has already existed for various different tasks to work for deep learning. A part of the challenge that we really focused on was, *How do you best create an infrastructure that facilitates this deep learning task?* There are a lot of small research centers. There are a lot of big research centers, which have a lot of heterogeneous hardware out there. All of it may not be best suited for machine learning. One of the best things that we did focus on for this research was, how do we best help people in those situations create deep learning stacks that would work really well at performing and best facilitate the learning task at hand.

Carson: I think this actually calls to mind a really interesting part of this project and something that was a great idea that the ETC data science group and the CERT data science group were able to come together on. So Ritwik faced this. *We have got terabytes of data.* It took a week to download. You were running up the bare metal. The electrons were causing you problems sometimes. From my perspective as a mathematician and in the CERT data science group, we mostly have a statistics background.

So, I don't know how computers work at all, but I know kind of a lot about how linear algebra works and how the theoretical gains that need to be addressed and made to do better on this project. It was actually really awesome to be able to realize that we needed both of those perspectives on the project. A group that is trying to do a project like this themselves needs to have people that have both of those skills. If your data is huge, you can't go without someone who knows about the electrons. Also, if your data is huge, you can't go without someone who knows about the math. You need both.

Ritwik: You can see that big groups like Google, Microsoft, Apple, do really well at this, right? They have people who are cloud engineers, who are really good at the infrastructure, and you have people who are just deep learning scientists, or machine learning scientists, who are really good at the math. They work together in one lab to make themselves the best AI labs in the world. Even if you are not at that skill, even if you are not at Google, even if you are just a three-person team, it is essential to be able to identify your underlying infrastructure limitations and identify workarounds to those problems so that both tasks can be best facilitated. There are tradeoffs that come with both of those, both on a learning side and on an infrastructure side. You have to identify the best tradeoff for your situation.

Will: So there are features to be discovered in what infrastructure is needed?

Ritwik: Right. Part of the research that we did focus on is, *What kind of tradeoffs would you have to make to do this?* Again, as Software Engineering Institute, we focus on the software engineering part of it. What engineering challenges are there to architect a system that would work on…This is a relatively small dataset, small being five terabytes, right? Because there are data sets that are much bigger than that. But even at five terabytes, how do you handle a dataset of that size? We focused heavily on that research.

Thank you for joining us. Links to resources mentioned in this podcast are available in our transcript. This podcast is available on the SEI website at sei.cmu.edu/podcasts and on Carnegie Mellon University's iTunes U site, as well as the SEI's YouTube channel. And as always, if you have questions, please do not hesitate to send us an email at info@sei.cmu.edu