Transfer Learning: Avoiding Regulatory Pitfalls in Medical Devices

With Prof. Dr. Oliver Haase, Prof. Dr. Christian Johner

Transcript

00:00:05 Speaker 1

Medical Device Insights, a podcast by the Jona Institute for medical device manufacturers, authorities and notified bodies.

00:00:18 Speaker 2

I welcome you to another episode of our podcasts and we will once again today enter the topic of machine learning, artificial intelligence.

00:00:29 Speaker 2

In a way, it's a deepening of a conversation I already had with Professor Oliver Haase, where we talked about the validation of machine learning libraries.

00:00:42 Speaker 2

For all those who were not yet part of this episode, I link it to you again below in the accompanying materials.

00:00:49 Speaker 2

For all those, Oliver, who don't know you yet, could you introduce yourself very briefly, please.

00:00:54 Speaker 1

Yes, very much, Christian.

00:00:56 Speaker 1

Thank you very much for letting me be there again.

00:00:59 Speaker 1

I'm a computer scientist, or more precisely, I have a professorship for software engineering and have been involved in software validation and software verification for many years.

00:01:11 Speaker 1

and More recently, I am applying these methods or these interests in particular to methods on machine learning methods and looking at how to use these methods, especially in regulated markets

00:01:26 Speaker 1

how the medical device market can actually use such machine learning methods.

00:01:32 Speaker 1

In other words, the key question that concerns me is how to check and prove the correctness of a machine learning application in a regulatory-correct and comprehensible way.

00:01:45 Speaker 2

Of course, this also makes you the ideal man to discuss a very, very demanding topic.

00:01:53 Speaker 2

namely the question of how we deal with transfer learning and that may not mean anything to everyone.

00:02:00 Speaker 2

That's why my initial question to you, what is transfer learning anyway?

00:02:05 Speaker 1

Yes, so transfer learning means that you use already trained, so-called pre-trained neural networks.

00:02:14 Speaker 1

So you mostly apply this to neural networks, that is, that you use such already pre-trained neural networks

00:02:22 Speaker 1

in order to then retrain them with new training data, reuse them and thus also reuse them for new tasks, which are usually classification tasks, or then apply them to new tasks.

00:02:36 Speaker 1

And at the moment, at least that's mainly done in the areas of image recognition and natural language processing.

00:02:46 Speaker 2

You just said that they are already pre-trained.

00:02:48 Speaker 2

Is the idea correct that the

00:02:52 Speaker 2

weights and the threshold values in these neural networks, since they are already preset, are preadjusted.

00:03:00 Speaker 1

Yes, exactly, that's exactly how it is.

00:03:02 Speaker 1

So these are networks that are already trained on other data and therefore already have weights that allow them to recognize certain things, to carry out classification tasks.

00:03:17 Speaker 1

and the whole thing makes sense for several reasons and is also possible that you do it.

00:03:24 Speaker 1

So that you use these pre-trained nets and retrain again.

00:03:29 Speaker 1

But there are, there are several reasons, and the first reason why this is done is that such deep neural networks, which are in question, are used for image recognition, for example,

00:03:43 Speaker 1

that they work so well above all because the layers of this network or networks recognize increasingly complex or abstract properties or features.

00:03:57 Speaker 1

This means that the foremost layers of such a network recognize features, such as point clusters, edges, and other similar or other geometric formations,

00:04:12 Speaker 1

the posterior layers then recognize increasingly complex properties, such as whether they see a building or dogs or cats or even conspicuous skin changes, for example.

00:04:24 Speaker 1

And because it is the case that these front layers recognize such concrete things, exactly these front layers are also very universally applicable for the most diverse types of images and for the most diverse types of tasks.

00:04:42 Speaker 2

Aha, that's exciting.

00:04:43 Speaker 2

This means that when we talk about pre-trained nets, it doesn't mean that somehow all weights are already half-adjusted and then fine-tuned at the end, so to speak, but you actually have a division between the lower layers or you just mentioned the front layers, where you leave the weights and thresholds the same and only afterwards then adjust the weights and thresholds in the back or upper layers to the special Case

00:05:09 Speaker 2

can you say that?

00:05:11 Speaker 1

Yes, you can, you can say that just as well.

00:05:13 Speaker 1

So I could go into a little more detail about that later, we'll talk about how to do this retraining, this retraining.

00:05:21 Speaker 1

And it is, that's exactly how you anticipated it.

00:05:25 Speaker 1

So man, you then hold the weights, the front weights in the front layers, you hold them tight.

00:05:31 Speaker 1

because you can recognize features with it.

00:05:34 Speaker 1

But this first training process was exactly good in this context, that these front layers can recognize the low-layer features well, i.e. can recognize the concrete features well.

00:05:47 Speaker 1

That's why you hold on to these weights and the ones in the back, you train them anew.

00:05:51 Speaker 2

O.

00:05:52 Speaker 2

K., but now I've interrupted you, because you've already started to mention the advantages of pre-trained models.

00:06:01 Speaker 2

you have already mentioned one disadvantage, namely that it can be a bit more demanding in terms of regulation.

00:06:06 Speaker 2

So let's get back to the advantages.

00:06:08 Speaker 1

Yes, yes, of course, these are, these are partly advantages and partly they are simply reasons why you, why you can do this, why you have this opportunity.

00:06:18 Speaker 1

So we have now seen as a first point that the front layers are well suited because they can do things that can be used quite universally.

00:06:30 Speaker 1

A second reason why this can be done and why it makes sense is that such modern neural networks are very, very large structures.

00:06:40 Speaker 1

For example, if you look at the different variants of Google's Inception architecture, which are state-of-the-art neural networks, these networks have dozens to hundreds of layers and tens of millions of weights.

00:06:57 Speaker 1

And such networks will then also be

00:07:00 Speaker 1

with tens of thousands, sometimes hundreds of thousands of images.

00:07:04 Speaker 1

For example, the image database Imagenet, which is used in science as a database for many benchmark tests, contains more than 14000000 labeled images.

00:07:16 Speaker 1

And of those 14000000, 1.2 million are used as training data in the

00:07:23 Speaker 1

ImageNet Challenge, which was carried out until 2017 and in the course of which increasingly complex networks were created.

00:07:32 Speaker 1

So it's about very, very much data with which these extremely large networks are trained, and that's why such training is also very resource-intensive.

00:07:44 Speaker 1

You need

00:07:46 Speaker 1

extremely powerful hardware.

00:07:48 Speaker 1

You consume an enormous amount of electricity and the whole thing simply takes a very long time.

00:07:53 Speaker 1

And now you can argue and rightly argue that for a powerful medical device, such resource-intensive training is quite acceptable, that you can expect that, that you just do it.

00:08:07 Speaker 1

And that's true, but there's another component to that.

00:08:12 Speaker 1

If a training is very complex, very time-consuming, very computationally intensive, then you have all the less leeway to experiment with different training parameters.

00:08:26 Speaker 1

Because in reality, it's not enough to take your neural network once and then train it on a lot of data and then it's done.

00:08:36 Speaker 1

Instead, you usually have to experiment with a lot of different training parameters.

00:08:42 Speaker 1

and of course, the more complex such a training process becomes, the more difficult it becomes to experiment with these many different parameters, because of course you always end up with a limited amount of resources to produce your medical devices.

00:08:58 Speaker 2

That is, if an auditor were to argue now, yes, for safety reasons it would be much better if they would train everything again from 0 to scratch, yes, so as not to somehow

00:09:11 Speaker 2

Including a black box would be a possible counterargument, if I understand you correctly, this can be exactly the opposite, namely because the models are pre-trained, we have many more possibilities to find the optimal set of hyperparameters.

00:09:26 Speaker 2

Can you argue that way?

00:09:28 Speaker 1

Yes, that's a very good argument, you can argue that way and then there are other aspects, such as

00:09:38 Speaker 1

the third reason why you can do this, plays a very good role in this argumentation.

00:09:45 Speaker 1

Namely, if you train from 0 on, i.e. from the scratch, with an empty network, then you need very, very much labeled data.

00:09:56 Speaker 1

I just talked about the fact that you have several million data in ImageNet, that you feed the Inception networks with 1.2 million data for training, for example

00:10:08 Speaker 1

And so much labeled data, you have it in special domains, such as in the health sector, you usually don't have it at all.

00:10:20 Speaker 1

You are dealing with comparatively small amounts of training data.

00:10:25 Speaker 1

This means that with this data, you could not train such enormous and such powerful networks, you could not train them at all.

00:10:33 Speaker 1

This is only possible

00:10:35 Speaker 1

by pre-training the networks on the basis of more general data and then giving them the finishing touches, so to speak, by giving them the specific medically relevant data.

00:10:48 Speaker 1

So if you wanted to train them from 0 to 0, then the alternative would not be to use equally powerful neural networks, but to switch to smaller networks.

00:11:01 Speaker 2

Which may then not be as efficient.

00:11:04 Speaker 1

Which may not be so powerful, although we are already in the discussion about how well this works and the state of science is still somewhat divided at the moment, but we can also go into it separately.

00:11:25 Speaker 2

Yes, then maybe let's say again how it works, this retraining of the network, so how you proceed and

00:11:34 Speaker 2

whether you just add additional training data and what kind of training data you can add again so that you can get from the pre-trained to the fully trained model.

00:11:44 Speaker 1

Yes, exactly, exactly.

00:11:47 Speaker 1

So it's like this, in the simplest cases, you take the existing network, give it additional training data, run it through it and then you're done.

00:11:59 Speaker 1

This is the simplest case and

00:12:01 Speaker 1

And unfortunately, it's usually not exactly like that, but the first adjustment that you usually have to make is that you replace the result output, the so-called output layer, and adapt it to the new task.

00:12:18 Speaker 1

For example, if you use a network pre-trained in the Imagenet, then this network can usually classify images you see into 1000 different categories.

00:12:32 Speaker 1

This includes many things of daily life, such as mashed potatoes, star anise or shower curtains.

00:12:40 Speaker 1

So the net saw one of these objects.

00:12:45 Speaker 1

And because these are 1000 different categories, the output layer of such a network consists of 1000 neurons, each of which represents a possible category.

00:12:59 Speaker 1

And if you now want to use such a net to no longer recognize star anise and shower curtains, but for example 20 different types of skin lesions,

00:13:11 Speaker 1

then the first thing to do is to replace this output layer with one that does not contain 1000 neurons, but only 20, of which each of these 20 neurons represents a different type of lesion.

00:13:26 Speaker 1

This is the first adjustment that you usually have to make.

00:13:30 Speaker 1

And the next adjustment or the next question that you have to answer for retraining is the one we discussed earlier, namely at what point.

00:13:42 Speaker 1

of the net or up to which point of the net you freeze your weights, i.e. do not change them, because you have exactly the knowledge that they have gained from the first training run, because you want to keep it and at what level you no longer do it and let weights train again.

00:14:02 Speaker 1

So that's this discussion from earlier, that the front layers recognize very concrete things and that this

00:14:09 Speaker 1

I want to use this terrain and that in the back layers you put these concrete things together to form more abstract concepts.

00:14:18 Speaker 1

So you have to draw such a separation somewhere in the net, draw such a dividing level and say, up to this point I'll freeze and from here on I'll train again.

00:14:29 Speaker 1

And where this line is drawn cannot generally be said in advance.

00:14:34 Speaker 1

You have to try it out, you have to try it out very carefully and see what results you get.

00:14:41 Speaker 1

And that brings us back to the question that we, or with the topic we also had earlier, that such a training process usually does not consist of letting it run once, but of dealing with different parameters.

00:14:56 Speaker 1

and experiment until you find the best configuration.

00:15:00 Speaker 2

Yes, that means that we actually have 2 dimensions in which we have to experiment, one level of hyper-parameters and the other time the layer, so to speak, from which you forget what you have learned so far and let it be retrained.

00:15:15 Speaker 1

Yes, exactly.

00:15:16 Speaker 2

How well does that work in practice?

00:15:19 Speaker 2

So, if you look at it perhaps right now with the example

00:15:22 Speaker 2

Continue to shimmy that you have already hinted at, namely medical image recognition, yes.

00:15:27 Speaker 1

Exactly, we have already touched on it very briefly, this is actually still a controversial question in science at the moment.

00:15:37 Speaker 1

There are results, there are positive results that show that it works well.

00:15:42 Speaker 1

So there is very, very specifically in the field of skin cancer detection, there are some scientific works and also actually in the already put into practice

00:15:56 Speaker 1

Insights by seeing that it works well by using exactly the aforementioned Google Inception networks and retraining them and showing that you can achieve pretty good results with them.

00:16:12 Speaker 1

So there are promising, promising things already now, and on the other hand, there is also research work,

00:16:22 Speaker 1

those who come to rather critical results.

00:16:25 Speaker 1

So there is a recent comparative study by Google that compared how well retrained networks work compared to smaller and specially trained networks and the

00:16:44 Speaker 1

who could not measure any significant differences.

00:16:47 Speaker 1

So, as is often the case in the early stages of research, it is not quite clear at the moment.

00:16:55 Speaker 1

But one thing is already clear, namely that it will play an increasingly important role in the future.

00:17:05 Speaker 1

Because it works all the better the more similar the domains you are dealing with.

00:17:11 Speaker 1

So if you have a pre-trained net

00:17:13 Speaker 1

from the field of medical image recognition and then applies it to new questions in the medical field, then of course it will work much better, even better than if the net was fed with photos from general life.

00:17:31 Speaker 1

And that's why it is to be expected that there will be more and more medically pre-trained networks in the future,

00:17:41 Speaker 1

which can then be applied to new questions by retraining them, and that is then the next level of reuse.

00:17:50 Speaker 1

So machine learning has a lot to do with reuse.

00:17:55 Speaker 1

You have predetermined algorithms, predetermined network architectures that you reuse and and the obvious next stage will be,

00:18:05 Speaker 1

that you already use predefined nets and then train them to your needs.

00:18:12 Speaker 2

These are probably all things that will change quickly over time, depending on how scientific and technological progress progresses here.

00:18:21 Speaker 2

And these are exactly the things that may make some auditors break out in a cold sweat.

00:18:28 Speaker 2

Yes, what can he actually expect now?

00:18:31 Speaker 2

How would you do this

00:18:33 Speaker 2

pre-trained models see this topic from a regulatory point of view.

00:18:37 Speaker 2

So maybe also related to this, how is this still a sub, yes so is this software or is it now data, how do I have to look at it regulatorily, what are your thoughts on this?

00:18:49 Speaker 1

Yes, that's actually not completely obvious, so I don't envy the auditor, the one who has to deal with it.

00:19:00 Speaker 1

Because, of course, this topic does not appear in the standards and harmonized norms for the time being.

00:19:08 Speaker 1

That's why it's worth taking a top-down look at it and finding out that the use of pre-trained models takes place as part of the software unit implementation,

00:19:27 Speaker 1

which is regulated according to the harmonized standard IEC 62 304.

00:19:31 Speaker 1

So that's something that happens within this phase, within the machine learning pipeline, which then makes up the software unit implementation in this case.

00:19:43 Speaker 1

And that in turn means that all these phases of the machine learning pipeline, in which in this particular case there is a pre-trained network at the beginning of this particular case,

00:19:56 Speaker 1

All these phases must be carried out and documented according to the state of the art, and this applies to data collection, data analysis, pre-processing, selection, configuration and training of the model and model evaluation.

00:20:14 Speaker 1

In short, the use of a pre-trained model is regulated by the fact that

00:20:25 Speaker 1

that it must be carried out and documented as part of the software unit implementation according to the state of the art.

00:20:34 Speaker 2

Yes, there we have, I would also like to refer again to the other podcast that we already had on the topic of Machine Learning Library and where you had already explained that we have to look at it from 2 perspectives.

00:20:47 Speaker 2

We have the perspective of the 62 304, especially now when it comes to the topic of sub and code development

00:20:55 Speaker 2

that becomes part of the medical device and the 13 485 that requires Computerized System Validation.

00:21:03 Speaker 2

That is, the validation of software that we need in this process, possibly also of the training, which does not become part of the medical device.

00:21:13 Speaker 2

Would you like to say something about this distinction again?

00:21:18 Speaker 1

Yes,

00:21:18 Speaker 1

Yes, of course, so thank you very much for picking it up again, Christian.

00:21:22 Speaker 1

That's how it is here in this case.

00:21:27 Speaker 1

Such a pre-trained network is first of all SOOP or is SOOP in any case, because this pre-trained network is usually or not only as a rule, but always

00:21:39 Speaker 1

is implemented in the form of a machine learning library.

00:21:43 Speaker 1

This is usually TensorFlow, Carers, PyTorch.

00:21:47 Speaker 1

In one of these technologies, these pre-trained networks are implemented and if they end up in the medical device, then they also become soup and must be validated as such.

00:22:01 Speaker 1

And as you mentioned, Christian, we had our own episode on that and also

00:22:09 Speaker 1

During the training process, these libraries and in this case even the pre-trained networks play the role of a software tool and then have to be validated according to ISO 13 485.

00:22:25 Speaker 1

And then there is now in this case

00:22:29 Speaker 1

requirements that go beyond this when we are dealing with pre-trained networks.

00:22:36 Speaker 1

Namely, there are still the questions to be answered as to why this chosen pre-trained model is actually well suited for the envisaged question.

00:22:47 Speaker 1

This must be justified.

00:22:49 Speaker 1

This must be justified, for example, on the basis of the model architecture used, on the basis of possible similarities between source and target files.

00:22:58 Speaker 1

And ideally again through comparative studies.

00:23:01 Speaker 1

And also, as we said earlier, it must be justified why weights were frozen at which point or the model was frozen, up to which point and from where the new learning was done.

00:23:17 Speaker 1

Here, too, it's not enough to say that we've done this up to layer XY and that's it, but you need a good justification.

00:23:25 Speaker 1

why it was done this way and comparative studies help there again.

00:23:29 Speaker 1

And that shows once again how important a well-founded evaluation of the trained model actually is in machine learning.

00:23:39 Speaker 1

and And in addition

00:23:43 Speaker 1

Is the use of statistical methods at least recommended, if not indispensable?

00:23:50 Speaker 2

Yes, even mandatory, because we even have an explicit chapter in the 13 485 ,n, namely on the subject of validation, where this is exactly what it is about, justification according to statistical methods.

00:24:03 Speaker 1

Yes, absolutely.

00:24:04 Speaker 1

So thank you very much for the hint.

00:24:05 Speaker 1

So this and exactly this requirement is, in my experience,

00:24:12 Speaker 1

in the area of model evaluation.

00:24:17 Speaker 1

So we rarely see in practice that models, we always see that they are evaluated, but we rarely see that this evaluation is then checked for its statistical significance.

00:24:33 Speaker 1

This is perhaps a topic to which we, to which we would like to dedicate a separate podcast here, because it is relevant

00:24:41 Speaker 1

even independently of pre-trained networks.

00:24:43 Speaker 2

Yes, maybe very briefly for the listeners, so that they know how to classify what this further contribution will be about.

00:24:50 Speaker 2

It is not enough to say afterwards that I have achieved significance or specificity or any other quality parameter, but we also need information.

00:25:00 Speaker 2

What is the probability that this parameter is in a certain confidence interval and that exactly at this information

00:25:09 Speaker 2

we usually see little more than any justifications and you have to go into it again, because otherwise the statement about significance says nothing at all about how much you can rely on this number.

00:25:24 Speaker 1

Yes, exactly, absolutely, and that's also because this way of thinking is not so widespread in the world of data scientists.

00:25:37 Speaker 1

So these

00:25:38 Speaker 1

these requirements, as you mentioned, from ISO 13 485, that you also have to back up your experiments with statistical significance.

00:25:50 Speaker 1

This is, this is not the way the world of thought and the way of working that has been done in machine learning or data science so far.

00:26:00 Speaker 1

And that‘s why there‘s a bit of a break between what the regulatory requirements for medical devices are.

00:26:08 Speaker 1

and, let‘s say, the state of the art in the field of machine learning.

00:26:13 Speaker 1

So I think there is an urgent need for improvement.

00:26:16 Speaker 2

Yes, on this occasion maybe also a note about the machine learning guide, which is currently also published by the W.

00:26:23 Speaker 2

H.

00:26:23 Speaker 2

O.

00:26:24 Speaker 2

is also developed.

00:26:25 Speaker 2

We have already put away a lot of these thoughts and I recommend it to any auditor or company that uses machine learning that wants to adapt its medical devices to

00:26:36 Speaker 2

especially because the notified bodies now also use this guideline in a slightly modified form than their own and also regularly use it during audits.

00:26:48 Speaker 2

Perhaps one of the most common violations of the guideline that we often come across and we have just hinted at this today, we often observe that the manufacturers

00:27:01 Speaker 2

train the models and simply say afterwards, yes, this significance or specificity or whatever that parameter is, that was what we achieved and that‘s their performance claim, so to speak.

00:27:12 Speaker 2

However, it would have to be correct that the intended purpose, the risk management, the state-of-the-art of the performance parameters are derived and models are then chosen that are able to meet these requirements.

00:27:28 Speaker 2

So sometimes, unfortunately,

00:27:30 Speaker 2

To observe a bottom-up approach instead of a top-down approach, and that's not quite clean from a regulatory point of view.

00:27:38 Speaker 1

Yes, exactly, exactly, and if I want to do it again very briefly, thank you very much for the good summary, if I may add something briefly, it should actually ideally be the case that the requirements for the forecast performance result from the intended purpose, that you then evaluate your model, measure a certain performance and

00:27:59 Speaker 1

and then, in a final step, proves that the measured performance also has a certain statistical significance.

00:28:08 Speaker 1

And this step is generally missing and that's what we can talk about in more detail in another podcast.

00:28:19 Speaker 2

Yes, that's already a perfect summary and an outlook in one sentence.

00:28:24 Speaker 2

all that remains for me to say is that, as always, we have included further links and materials in the accompanying materials, including an article specifically or even 2 articles specifically on the topic of validation of machine learning models and validation of pre-trained networks.

00:28:45 Speaker 2

I think that will be very exciting for the listeners again and these thoughts, we will weave them into this guide again.

00:28:53 Speaker 2

And.

00:28:54 Speaker 1

If I may allow myself a brief addition, Christian.

00:28:57 Speaker 1

We have, we recently just released the Inception V.

00:29:03 Speaker 1

4 model from Google completely validated, i.e. 1 of the largest currently available neural networks and which, as I mentioned at the beginning, is already being used successfully in the medical field and has been retrained, and have thus shown on the one hand,

00:29:23 Speaker 1

that the ways in which we do this, that our method of validating machine learning models in accordance with the law, that we can also apply it well to very, very large networks, and on the other hand, we have also created a basis for a special kind of pre-trained network, which may be useful for one or the other

00:29:51 Speaker 1

handset is actually used or could be of interest.

00:29:56 Speaker 2

And I would recommend all of them to contact Professor Haase.

00:30:00 Speaker 2

This helps to answer exactly such questions or, in the special case, to validate these machine learning libraries, including those in front of the trained networks.

00:30:13 Speaker 2

And that's what we all do.

00:30:15 Speaker 2

manufacturers who use these processes.

00:30:18 Speaker 2

So, his contact details can also be found below in the accompanying materials.

00:30:23 Speaker 2

Oliver, thank you very much for these exciting insights.

00:30:28 Speaker 1

Yes, thank you very much, Christian too, was a pleasure.

00:30:31 Speaker 1

I was happy to be here for the second time and I hope we almost have something like a tradition now.

00:30:36 Speaker 2

Exactly, and the topic for the number 3 has already been set today, I'm looking forward to it.

00:30:41 Speaker 2

Oliver, thank you again.