Daniel Reukert July 2018 SA - DEPTH - 12 MINS
[MALE RESPONDENT]
[Other comments:]

Now for the next talk, much awaited at this point, with comments before, Daniel Reukert from the Department of Computing at Imperial College will tell us about machine learning and artificial intelligence as applied to UK Biobank data.

Thank you very much for the introduction and thank you very much for the invitation to talk to you. I'm probably a more outsider here, I'm a computer scientist, and I wanted to highlight to you a bit about what we can actually do with using AI to help with the analysis of UK Biobank. I wanted to start with a comment which you have probably read in the press, from Geoff Hinton, who's one of the founders of modern machine learning, who said, actually, we should really stop training radiologists now. That, of course, in the context of Biobank, is an issue, because actually there are not enough radiologists around. Steffan has already highlighted what a huge effort it is to actually analyse imaging data automatically, but this statement is perhaps a bit over the top. So perhaps a slightly better statement is from another colleague who said, actually, will AI replace radiologists? The answer is no, but radiologists who do AI will probably replace radiologists who don't. I think this is actually true, perhaps we can also say about UK Biobank scientist. Perhaps those UK Biobank scientists who use AI will replace those who don't use AI. I wanted to highlight two specific aspects where AI could be actually useful.

Steffan already has pointed out something about how we derive image phenotypes from the data, but I also wanted to highlight how we can use machine learning for automated quality control. In our group, we have developed, we have been working machine learning techniques for quite a while. This is a machine learning technique called deep learning, which has been very popular in imaging, in computer vision, and can be used to analyse medical images. Basically what you do is, you put an image in, and you get, for example, in this particular case, a semantic labelling of the image out, for example, into the structures of the heart. Now, all of these techniques usually require a lot of training data to work reliable, and this is one of the great opportunities in the collaboration with Steffan's group here in London, but also his colleagues in Oxford. They had actually annotated already 5000 subjects, and 5000 subjects really consists of almost 100,000 2D images. That's actually a very large amount of training data. What we did is, we took this training data and tried to see, can we actually tell the machine how to analyse the remaining images in UK Biobank, so that we can actually upscale the analysis to the whole 100,000 eventually? Here is what you actually see, what is the outcome of this sort of training the algorithm with those segmentations.

You can now actually fully automatically analyse these [unclear word 0:03:28.9] sequences, not only the [?an-systolic and diastolic time point 0:03:32.0] as was already pointed out, but fully automatically in the entire time series, and you can now derive measurements from this time series. We have really done quite a lot of evaluation together with Steffan's group to find out, how well can the computer do this? Because that's of course the ultimate question we have. It turns out that if we do automatic measurements and we compare that

to the consensus of multiple observers, that is actually in the same range as a disagreement between multiple observers. So the conclusion is that the computer doesn't have superhuman performance, but actually, the computer does a job as well as another cardiologist or radiologist would do. Really, we are within the inter-observer variability, which is what we can actually hope to get to because, of course, our algorithms are trained with data from humans. We can't really do the job better than the humans can.

We can now derive on the entire 20,000 - I mean, we've done this now for 20,000 subjects. We haven't yet done this for 25,000. We can derive all these different numbers. We can start to look at, for example, how different cardiac image-derived phenotypes are related to age, we can look at left ventricular, right ventricular volume, atrial volume and so on. We can correlate it to age, and all sorts of other measurements. One of the things which is very interesting is, we're now also at the position where we can actually do what Paul already showed you for brain imaging-derived phenotypes. We can now look at, for example, how do cardiac phenotypes correlate to different measurements in Biobank related to lifestyle and other things? What you can see is, you can already see a number of different observations. For example, of course cardiovascular health is very closely related to physical activity score, and you can actually now mine the data in a very interesting way. And hopefully we'll be able to, as new data comes into Biobank, apply these image analysis techniques to new data and Biobank.

I also wanted to highlight a different application, which probably we hadn't initially anticipated for machine learning or AI, namely, can we actually do automatic quality control with the data in UK Biobank? One of the things you are probably not that familiar with, is that the image acquisition is not only a technically challenging process, but also requires really well-trained operators, and requires quite a lot of careful manual input in running these sequences. What we have done is, we have used machine learning to train based on data from Biobank, whether for example the right amount of the heart has been covered, whether there is patient motion between different slices, whether their contrast is as we expected. For example, here we automatically calculate whether the operators when they planned the scanning have actually covered the entire left ventricle. Because of course, that's very important. If you miss out part of the anatomy during the imaging, you actually don't really have a complete data set. We wanted to automate that in order to identify any problems. One of the results you can see here, we've for example applied this to 20,000 subjects again in Biobank, and you can see that for example in the first year of Biobank, for example, many of the subjects which were scanned had only a coverage of around 90 per cent to 100 per cent of the heart.

I think this was already identified using Biobank's own quality control procedures. They then subsequently updated the way they actually acquire the data, or the instructions they give to the operator. And what you can see is, in subsequent years, actually the coverage has significantly improved. What we can do is, now we can automatically detect trends in the data, and insure that actually this might be used to, for example, update guidelines which are given to the operators. I also wanted to finally highlight how AI might transform how we do the image interpretation in Biobank, and how we actually correlate image-derived phenotypes to outcome. At the moment we all follow this protocol. We derive some features of phenotypes, for example, left

ventricular volume, ventricular volume in the brain, we look at structures in the brain, measure them. We then build a model to try to interpret this, and correlate it to some measure of interest. But really the introduction of deep learning techniques allows us, actually, to bypass these two separate processes.

Because we're not really sure that we're extracting the right features for looking at what we actually want to study in terms of a predictive model. Deep learning offers us the opportunity to end-to-end train this pipeline, and therefore be much more powerful in developing predictive models. For example, for age we can also in these predictive models easily integrate other types of information and potentially really come up with a model which is predictive of whether a subject will develop disease. Now, of course, there's also a downside to this. This requires us also to work on techniques which we can use afterwards to explain what we have found, and interpret the features or the information we've modelled with AI. There is no use, really, just having a black box approach to this, but I think this is very exciting. Finally, again, something which Biobank probably hadn't really intended when it was set up, is that we can also really use the data to develop better ways of acquiring the images in the first place. What you see here, on the left-hand side, is for example an image which is acquired six times faster than the normal acquisition, and therefore much easier for the patient to tolerate, and then using data in UK Biobank, we can actually learn how to recover the full information which is still in this data set, in order to actually reconstruct an image which is useful for clinical interpretation.

This is actually really the significant advantage of having so much data to mine with AI, and here you see how we can use a similar approach to learn how to reconstruct high-resolution data of, for example, the heart, from low-resolution data. Again, Biobank is a fantastic resource for researchers in this community to develop novel and innovative solutions for this. Finally, really, I hope that at the end we will be able to develop solutions which can really do end-to-end optimisation of the entire process of acquiring data, reconstructing the data, and interpreting the data using machine learning. Of course, data sets such as UK Biobank are absolutely critical in enabling such research efforts. We have for example at the moment, a programme grant funded by EPSRC, to do exactly that in looking at images of the heart. With that I want to stop here and thank you very much for your attention. Thanks.

[END OF TRANSCRIPT]