COVID Information Commons (CIC) Research Lightning Talk

Transcript of a Presentation by Michael Pazzani and Albert Hsiao, (UC San Diego), September 28, 2020

Title: RAPID: Explainable Machine Learning for Analysis of COVID-19 Chest CT

Michael Pazzani Database profile

NIH Project #: 2026809

YouTube Recording with Slides

September 2020 CIC Webinar Information

Transcript Editor: Lara Azar

Transcript

Katie Naum:

Next, we have Michela Pazzani and Albert Hsiao of the University of California at San Diego who will be telling us about explaining machine management for COVID-19 management, so I'll let you guys take it away. Thank you.

Slide 1

Michael Pazzani:

Okay so first I'd like to thank NSF, particularly information intelligence systems, for funding this, and this is a collaboration between myself, a computer scientist and Albert, a radiologist with a deep background in deep learning, and we're using machine learning methods, particularly deep learning, to analyze CTs or x-rays to manage COVID-19. While a diagnosis is important, we're also concerned with understanding the severity of this disease from the imaging, and finally it's not sufficient just to say you have a 96% chance of having COVID but, or even to highlight a portion of the lower lung, but we aspire to label the images with things like, there's ground glass in the lower left lung. We're evaluating a variety of existing approaches for both classification and explanation and starting to develop new ones as well. The ultimate goal is for machine learning to acquire diagnostic signs that can be communicated to people, such as a clinician when they're doing diagnosis, or perhaps to teach peers or residents without even [*inaudible*] computers marking up the images. On the next slide we show the process of deep learning-

Slide 2

-and essentially we have a database, an existing database of images as well as new images that unfortunately we've collected at UCSD over the past few months of patients with COVID, and the goal is to take images of normal patients, those with COVID-19, those with other conditions and come up with a learning method that distinguishes them, and also provides the explanation. We're quite fortunate in that we get to leverage a lot of existing infrastructure that was already in place of doing imaging, sending it to the cloud for analysis, and then sending it to the clinic where the physician end user can observe it. All of this was in place already by Albert at UCSD Health and we've just had to modify the cloud diagnostic procedures with new data from COVID-19.

Slide 3

And from here I'll let Albert take over and explain a little bit more about how we're doing this.

<u>Albert Hsiao:</u>

Thanks Mike, thank you again for giving us the opportunity to present this work. It's definitely very technical in nature but also very clinical and impactful immediately as we're already using it in our clinic. The primary concept is really to develop AI algorithms that allow us to localize pneumonia. This is work that we started even before the COVID-19 pandemic began, but has been accelerated a lot because of the need. One really key important aspect of COVID-19 is that not every patient develops pneumonia, some patients do and some patients don't, some become asymptomatic, of course, but those that do, the severity of pneumonia on x-ray or CT provides us very good prognostic information and a lot of data is starting to come out with that. We've taken a very different strategy towards this U-net type segmentation approach as opposed to a lot of classification approaches that have been previously used, although both are feasible, and you can generate these kinds of probability maps and what-not - activation maps from classification approaches which we'll be exploring as well. The important aspect is quantifying the severity of illness, I guess, essentially gives us prognostic information, because ultimately we want to know which patients require hospitalization, which ones can stay at home, which ones require mechanical ventilation and which ones are likely to survive or not. And some of our initial data here is showing us that those patients with high likelihood predicted by the - by the algorithm are also the ones that tend to not survive and are the ones that tend to require intubation. So, this will give us really good data for how to best manage these patients so that's a really critical element of how-

Slide 4

-how this will all come into play in our clinic and hopefully many others through our collaborations. Our current results in COVID-19, this is one example of a patient with COVID-19 who presented to our clinic. Our AI algorithm produced our initial AI algorithm produced this result, very subtle that doesn't really highlight the areas of pneumonia that well as it was trained on initially only public data before COVID-19 and we came up with a strategy that uses active learning transfer learning to specifically identify good cases for us to train on applying transfer learning to that neural network. Using also concurrently performed CTs that were done to give us a better ground truth and that

that's given us a higher performance both on the internal and external data set and really highlights the pneumonia better so our initial algorithms that were in place we're replacing with these updated algorithms-

Slide 5

-thanks to the support of the NSF in this project we've been deploying this it's in our clinic and there's certainly articles online about it as well as a peer-reviewed publication that we brought out at around the time that we were investigating this initially. And our next steps are really to aggregate large data sets across multiple institutions, have multiple readers annotate markup the areas of pneumonia to sort of solidify the ground truth a little bit, use the CT as well and develop this comparable algorithm for CT in in the process. And ultimately, we want this algorithm to be explainable, is really critical for us to be able to use it clinically is to be certain that we're relying on features that really matter not, not sort of accessory features that the neural network sort of coincidentally saw associated with COVID but actually things are related to it and then. And then assess its clinical utility both in the detection of disease, distinguishing between other diseases that are quite similar like pulmonary edema, and then give us best management practices for these patients, so that's kind of where we're going. Thank you.