Closing the workflow loop: HistoQC for digital slides

Valerie Neff Newitt

July 2020—Unveiled in 2018, HistoQC, an open-source quality control tool for digital pathology slides, was an "awakening to a problem" and the kickoff of a conversation, says Andrew Janowczyk, PhD, its main investigator. And while it's hard to measure the tool's use because it's freely available, he says, reception has been strong. "It's been a pretty good ride," he says of its first two years.

Dr. Janowczyk, assistant research professor, Department of Biomedical Engineering, Case Western Reserve University, and colleagues designed HistoQC to make it easier and faster to identify and delineate artifacts and batch effects during routine slide preparation and digitization. "Manual review of glass slides and digital slides is laborious, qualitative, and subject to intra- and inter-reader variability," they wrote, stressing the need for an automated way to spot slides that need to be remade and regions that should be avoided during computational analysis (Janowczyk A, et al. *JCO Clin Cancer Inform.* 2019;3. doi:10.1200/CCI.18.00157).

Their solution, HistoQC, fills "an intuition gap" in artificial intelligence, Dr. Janowczyk tells CAP TODAY.

Pathologists train themselves over years, he says, to read through slides of suboptimal quality, if necessary, "because they're highly skilled, very well educated, and have learned to overcome these types of hurdles."

Dr.Janowczyk

"Digital technologies like artificial intelligence and machine learning are unfortunately not currently that skilled. They're not that robust in the presence of artifacts on slides," he says. "So something needs to be in place that guarantees that whatever uses a slide next, be it human or machine, knows what it is getting and can feel confident that it is exactly what it thinks it is. In other words, we want to make sure that before a sophisticated, yet relatively ignorant, algorithm is used, there will be a step to make sure that the sample it is to be applied to is appropriate."

"HistoQC quickly and efficiently does exactly that," he says.

Dr. Janowczyk, who is also a member of the cancer imaging program at Case Comprehensive Cancer Center, says the user interface was built so that HistoQC can be used on a regular internet browser.

A tab-separated value file in which image metrics are saved can be loaded into the browser-based front-end or any type of statistical toolbox, such as Excel, allowing any type of analysis to be performed, Dr. Janowczyk explains. "You simply double click the file on the hard drive of your computer, it will open up a web browser, and you can see results there immediately. From there, the user can go in and manipulate the user interface, which shows all the metrics and thumbnail images in real time." Now the user can look at, say, 1,000 slides in 15 to 20 minutes by scrolling through them. "This one's good. This one's bad. It becomes a very efficient process."

For clinical use, he says, as soon as a slide is scanned, it can be run through HistoQC and the quality can be evaluated before the pathologist sees it. "We can determine much sooner in the workflow if a slide is of bad quality. It doesn't have to go to the pathologist to get rejected," which means that "the large feedback loop becomes very, very small." Work has begun at a number of hospitals in this regard, he says. "The limitation is that you need to have a digital clinical workflow in place to take advantage of these types of digital tools."

Further, Dr. Janowczyk says security is built into the design. “There’s no connection to the internet whatsoever, so you can use it with nonanonymized, confidential patient data. Everything is self-contained.”

A comparison of HistoQC against manual QC by two pathologists on 450 images revealed an average agreement of more than 95 percent.

HistoQC got its start when Dr. Janowczyk had an idea for an unrelated study and turned to the repository of slides available through The Cancer Genome Atlas of the National Cancer Institute. “TCGA has about 30,000 slides, all free and publicly available and accompanied by a lot of data. It’s a very rich resource,” he notes. He found about 600 slides he could use for the study he wanted to do. “I was thrilled because I didn’t have to find new patients. These patients already exist, and they’ve already provided ethics approval.”

When he downloaded the 600 slides and started to look at them, Dr. Janowczyk could see that many were of lower quality and unsuitable for the study. “These slides were never really intended for analysis with computer algorithms. The cohort wasn’t designed for you to compute directly on the slide itself. They built it to the standard for human pathologists, who are robust to quality control problems.”

About 10 percent of the slides were not suitable for computer analysis, “which isn’t bad,” he says. “That’s about the norm.”

“But if it takes just one minute to look at 600 slides, you’ve already lost 10 hours. Right then I realized we need a way to look at slides more efficiently and identify parts of slides that are good and bad quality. Sometimes half of the slide may be bad but the other half acceptable,” he says.

In years past, when data sets were smaller, quality controlling of the images could be done manually, he says. With digital scanners increasingly prevalent, larger repositories are being constructed, resulting in more digital slides at a faster rate than ever before, so “it’s just not feasible anymore.”

“These same large cohorts are where the statistical power for our studies comes from. That’s where we’ll be able to identify small, nuanced biological signal. So we must be well positioned to take advantage of this increasingly massive amount of information.”

It wasn’t Dr. Janowczyk’s first attempt at using a larger cohort. He and colleagues had built a series of in-house scripts and filters over the years to identify artifacts and regions that were out of focus in slides they wanted to use. “Building HistoQC itself was a way to organize our experience into a single consolidated pipeline, as well as make it robust and easy to use for other people.” He says it can be used with little training.

“That’s how HistoQC began. I just sat down and started developing it,” he says.

Work began in 2017, and HistoQC was officially released in June 2018 at the European Congress on Digital Pathology, where it won an innovation award. The source code is available in an open-source github repository (github.com/choosehappy/HistoQC). “Literally anyone can download it and start using it. And because it is part of the open-source paradigm, other people can now contribute to the effort. It has become a living, breathing project in the sense that people we don’t even know can add new components. They can integrate it, send us the code, and we can merge it into the central public repository.”

Dr. Janowczyk says it is hard to know how many HistoQC users there are. “One of the issues with giving something away free is that you have no real way to track how many people have used it. That said, I can say we receive comments from users all the time. Institutions, too, are starting to use it. In Switzerland we are already seeing uptake in a number of hospitals, and in Scandinavia too.” HistoQC was built in collaboration with Michael Feldman, MD, PhD, of the University of Pennsylvania Department of Pathology and Laboratory Medicine, “and he is interested in deploying it clinically there as well,” Dr. Janowczyk says.

“HistoQC is the first available tool of its kind and it happens to be free, so it is kind of growing like wildfire.”

The HistoQC user interface showing (top) list of slides and their properties, (middle) plot of the different quality metrics to help identify outliers and batch effects, and (bottom) original slides next to fuchsia overlays wherein artifact-free regions identified by HistoQC are indicated.

Use of HistoQC will grow as users discover even more of its advantages, he says. There will be savings, for example, when staff knows precisely how long a stain is viable. “Stains are expensive. They are like oil in your car; you have to change them. But how often? Manufacturers will give you a conservative boundary to be sure the stain is working properly.” But with HistoQC, the quality of the stain can be measured over a period. “You will have a quantifiable measure that shows you if a stain is still good. You may find that you don’t have to change the stain as often as you thought.

“On the other hand,” he continues, “if something goes wrong with your stain, you would know immediately, as soon as those first few slides come out of the scanner. You can hit that big red stop button before you unintentionally create a thousand bad slides.”

Alternatively, at the end of a month a laboratory could see that last month’s slides were of lower quality and know not to use them in the future for computational analysis or potentially clear them out of storage to make room for better quality data. It gives pathology a QC paradigm that’s everywhere, he says—building cars, steel beams, buildings—but was not previously possible in pathology because it wasn’t a digital science.

In research, HistoQC may act as a proving ground for algorithms in development and in use. “Researchers need to know they are testing new algorithms on slides that meet specifications. The question is always, ‘How can we trust this algorithm?’” Dr. Janowczyk says. “Previously we tried to show trustworthiness by curating large collections of slides from different hospitals and seeing if the algorithm works consistently across those cohorts. That’s a brute-force approach to validating an algorithm. But with HistoQC we can computationally look at a million slides and identify 5,000 of them that are the most diverse and then test an algorithm on those.” If it performs well, he says, they know they have covered the spectrum of what they might expect to see in the real world.

When an algorithm is introduced at a hospital, the metrics on that hospital’s slides can be compared against the metrics on the slides used to develop the algorithm. “We will be able to say, ‘Your slides are very dark or light compared with the slides we’ve built our algorithm on, so it may not work as expected. You should proceed with caution.’ It’s an opportunity to raise red flags to make sure an algorithm is not used in a way that was not intended.”

HistoQC’s collaborators, in addition to Dr. Feldman, are Ren Zuo of Case Western, Hannah Gilmore, MD, of University Hospitals Cleveland Medical Center, and Anant Madabhushi, PhD, of Case Western and Louis Stokes

Cleveland VA Medical Center. They offered HistoQC open source to spur a conversation, Dr. Janowczyk says.

“Digital pathology remains like a wild frontier,” he says. “Everyone is still developing their own tools and workflows. So it was important for us to take a step back and think how we can create a way that people can work together, to make this a little more efficient, a little more formalized, and a little less chaotic.”

It’s not that QC wasn’t performed but that it wasn’t centralized and uniform, he says. “It wasn’t organized and reproducible in a way that another lab would be able to do it in exactly the same way and get the exact same results.” He views that as “bad science.”

“One component of this project was the intent to formalize a quality control process and put it out there for others to evaluate and discuss. In its own way, it’s a bit revolutionary. But we won’t know how much of an impact this will have until we look back in a few years. Hopefully by then new standards and regulations will have been created as a result of realizations derived from using it.”

Perhaps scanner manufacturers will build these types of tools into their systems, he says, so slides can be identified as poor quality while still in the scanner. “And some are starting to do that.”

Dr. Janowczyk doesn’t necessarily see HistoQC still being needed decades from now. “To be honest, I would be surprised if it was still needed. I think HistoQC represents more of an awakening to a problem. We’re pointing to that problem and saying, ‘Here is our perceived solution. Here is a strong beta prototype that everyone can use. And now we can start from there to have useful discussions about what we think quality control should be, and why and how it should be implemented going forward.’”

Valerie Neff Newitt is a writer in Audubon, Pa.