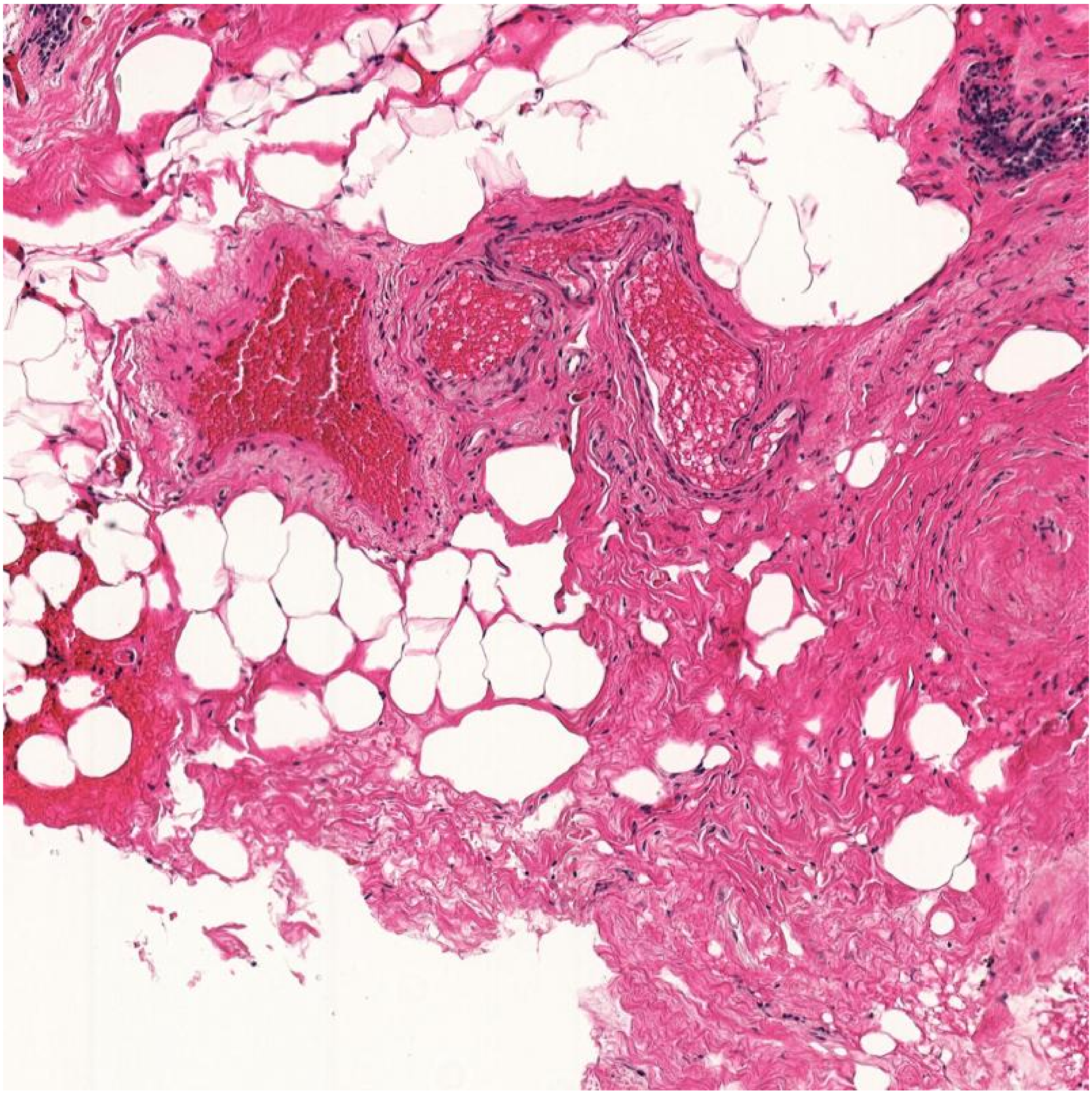
Doctors to get better access to digital data

July 9 2015



UA researchers are developing software that lets doctors more quickly and

effectively analyze digitized biomedical images like this one of breast tissue.
Credit: Ali Bilgin

The National Institutes of Health has awarded a $1.3 million grant to researchers at the University of Arizona to develop open-source software that will enable health care professionals and scientists to manage biomedical big data in digital form.

The advanced data compression software for the first time puts digitized biomedical data in a format and size that doctors, pathologists and other health care workers with limited resources and in remote locations will be able to access, analyze and store. Usable digitized data means quicker second opinions and diagnoses for patients.

"Advances in image compression technology for biomedical big data are essential to advance biomedical diagnostics and research and to save more lives," said Ali Bilgin, UA assistant professor in the departments of Electrical and Computer Engineering and Biomedical Engineering and principal investigator of the project. "We are in the middle of a global transition to digitization of biomedical data, and there's a lot of it out there—but in files too large to be transmitted, stored or retrieved."

Samples from a single patient over a lifetime can add up to a terabyte of data—enough to fill an entire hard disk, a huge amount of data that most pathology labs cannot handle.

The award is one of the first of 15 research projects funded through the NIH Big Data to Knowledge, or BD2K, Initiative, the agency announced in June. The NIH identified data compression as one of the initiative's key focus areas and gave the UA project ("Development of Software and Analysis Methods for Biomedical Big Data in Targeted Areas of

High Need") an impact factor of 11, the agency's second-highest possible ranking.

"Data compression software has the potential to address some of the most significant data storage, computing and sharing challenges facing biomedical researchers," the NIH said in its announcement.

Task-Specific Data Compression

The UA-developed software will compress the size of slide images up to 100 times without losing any detail or resolution from the original slide images. Bilgin's team is working with previously scanned digital images of slide samples from breast cancer patients, but the software will be applicable for a broad range of diseases.

"The medical community has been working to digitize biomedical data for some years, but to date the data has had little real effect on health care," Bilgin said. "Most labs are not equipped to receive or use the millions of gigabytes of information from data such as DNA sequence data or protein structure data. Our software will provide access to these files despite their huge size."

Perhaps most revolutionary, the new software will compress the digital images in specific ways for specific tasks.

"A pathologist may want to determine if an image indicates cancer. A research scientist might be more interested in identifying different categories of cancer cells in the sample," Bilgin explained. "Our technology will tailor the size and format of the digital image that is transmitted so a user doesn't receive a whole lot of data they don't need."

It should be a first, he said.

"Although it has been long recognized that image quality should be task-based, we know of no previous attempts to compress images for specific tasks," Bilgin said. "This is an entirely new way to think about image compression."

Engineering Better Biomedical Technology

Image compression was the subject of Bilgin's dissertation at the UA, from which he received his doctorate in electrical and computer engineering in 2002. Although he started out as an electrical engineer, he said, "I quickly realized that all of these skills I'd gained as an engineer could be applicable to problems in medicine."

Michael Marcellin was his graduate adviser and is co-investigator of the NIH-funded study. Marcellin, a Regents' Professor in the Department of Electrical and Computer Engineering and the UA College of Optical Sciences, was a major contributor to JPEG2000, the standard image compression coding system used today and the one that will be used in the biomedical image compression research.

Other participating UA researchers include Elizabeth A. Krupinski, professor and vice chair of medical imaging and associate director of the Arizona Telemedicine Program, and Amit Ashok, assistant professor of optical sciences and electrical and computer engineering. Researchers at Ohio State University, including Metin Gurcan, associate professor of biomedical informatics, also are participating.

Krupinski said the research has major implications for telemedicine, which uses electronic communications to transmit medical information, often to remote locations, to improve patient health.

"The use of telemedicine and the volume of associated digital images in health care are expanding exponentially," she said. "Many of these images, such as whole-slide pathology images, are extremely large and difficult for users to navigate through. Dr. Bilgin's task-specific compression techniques will make navigation more efficient and diagnostic interpretation more effective."

Second Opinions in Seconds

When a patient requests a second opinion today, pathologists send the patient's slide sample to another laboratory to prepare bioassays for the second doctor. The process can take days. The UA-developed software will transmit review-ready data in minutes or seconds, allowing users to view digital images and share information about them over vast distances in real time.

"This will help patients more quickly get second opinions and reduce the chances of diagnostic errors," Bilgin said.

It also will advance biomedical research and discovery.

"Let's say a pathologist sees a large cluster of a certain type of cell in their digital image," Bilgin said. "They can use the new software to request all digital pathology samples with a similar cellular feature. This has tremendous potential for increasing our ability to quickly identify disease."

Provided by University of Arizona