Will machines be able to judge a patient's prognosis? This prototype microscope aims to do part of the job.

SOFTWARE

The computer will see you now

From image–analysis software to lens–free microscopes that fit on a mobile phone, new tools are providing pathologists with clearer and more informative images.

BY KATHERINE BOURZAC

In the seventeenth century, natural historians such as Galileo, Antonie van Leeuwenhoek and Robert Hooke learned to grind lenses and make the first microscopes, revealing the hidden landscapes of life. They saw for the first time the cells in cork, blood and other tissues, and van Leeuwenhoek found swimming 'animalcules' in dental plaque and observed the movement of sperm.

Physicists and engineers are now trying to bring about a similar shift in perspective for microscopy. In most pathology labs, doctors diagnose diseases by poring over tissue slices on glass microscope slides — classifying tumours, for example, based on subtle visual cues that are difficult to quantify. But this is starting to change. Just as lenses once revealed vistas that were previously invisible to the human eye, so software is opening up a new window on biology.

The latest digital tools make it possible to do a visual search in microscopy images, automate diagnosis, and sync image data with the genomic profiles of tumours. Some researchers are even doing away with lenses altogether, creating computational microscopes based on inexpensive hardware that could be used for point-of-care diagnostics, particularly in poor areas with few doctors.

BIG DATA

Pathology has remained stubbornly analogue and qualitative, however. The experienced pathologist's main tools are glass slides, a compound microscope whose design has hardly changed in more than 200 years, and eyes that have seen thousands of tumours. "Most of a pathologist's medical decisions are based on morphology," the structural details of cells and tissues revealed under a microscope, says David Rimm, a pathologist at the Yale School of Medicine in Connecticut.

Just because a method is old is no reason to abandon it, of course. But advocates of digital pathology worry about inconsistencies that can lead to false negatives and misdiagnoses. Experienced pathologists are better than younger ones at identifying rare tumours, but they often disagree with one another and even with their own assessment of the same sample from weeks before.

One hurdle to digitizing clinical microscopy is the size and complexity of the images, says Metin Gurcan, who specializes in biomedical informatics at Ohio State University and was an early advocate of digital pathology. First, a biopsy is sliced into sections and placed on multiple slides. A digital image of a single slide, magnified under the microscope, has about 10 billion pixels and requires about 30 gigabytes of memory. A typical prostate biopsy, for example, uses more than 20 slides and needs about 600 gigabytes.

That's a lot of information for pathologists to scan through — and a lot of data for software to sift. "The number and type of cells found

in these images is mind-boggling," Gurcan says. One way to deal with this complexity is to use software that learns to recognize things in images the same way people do, but faster and more consistently.

VISUAL LEARNERS

Just as people learn by seeing many examples, so can software. In 2011, Harvard Medical School pathologist Andrew Beck built a tool called C-Path (for Computational Pathologist) by feeding learning software with images of breast-cancer biopsies from 248 patients, along with survival data[1]. The software learned to grade the severity of breast cancer and predict patient survival.

A human pathologist who looks at these biopsies under the microscope relies primarily on three features specific to cancer cells to decide how aggressive the tumour is. Do the cell nuclei have an unusual shape? Are the cells dividing? And are the cells connecting with one another as normal, or are they isolated? Pathologists qualitatively score each of these features to determine the tumour grade, a description of how aggressive the tumour is.

The C-Path system works by segmenting images into small regions called 'superpixels'. It identifies cell nuclei and cytoplasm within each superpixel, and compares the qualities of each superpixel — such as colour, texture, size and shape — with those of its neighbours. For breast cancer, this comparative analysis generates features related to both a sample's global structure and its fine-scale details, such as the average distance between the nuclei of cancer cells and normal cells.

After crunching the training set of images, C-Path came up with 6,642 features, describing not only the tumour cells themselves, which human pathologists focus on, but also the surrounding connective tissue, called the stroma. Indeed, Beck found that the morphology of the stroma was a better predictor of survival than that of the cancer cells alone: an area of stroma that was uniform was associated with a good prognosis, whereas stroma that was infiltrated by epithelial cells indicated more aggressive cancer. Based on its analysis of thousands of features, C-Path was able to predict patient survival more accurately than standard pathological analysis. Beck is now training the software on a broader range of samples, including images of whole slides, and normal breast tissue samples.

It is possible that highly experienced pathologists also look for some of the thousands of features spotted by C-Path but just can't describe them in words. Rimm compares the experience of spotting a tumour with recognizing your uncle in a photo. You can't articulate exactly how you know he's your uncle — is it his nose, eyes, clothing? You just know it's him. But the computer can quantify features in an image, and the analysis is repeatable.

Richard Levenson, a pathologist at the University of California, Davis, says that software such as C-Path has the potential to replace pathologists in assigning grades to tumours. Others believe that the right place for software is as an aide to help physicians navigate large digital images in real time — a second set of very sharp, superhuman eyes. Ulysses Balis, a specialist in pathology informatics at the University of Michigan in Ann Arbor, is developing this latter kind of tool: an all-purpose visual search program called SVIQ.

Balis demonstrates SVIQ with a digital image of a slice of colon adenocarcinoma. He can bring up different fields of the image and zoom in and out. If he finds something interesting — such as a cell that appears to be dividing — and wants to see if there are other similar features in the image, he clicks a 'scan' button and the software highlights all the parts of the image that look similar[2]. In this case, all the dividing cells turn red. The concept is similar to smartphone apps such as Google Goggles or
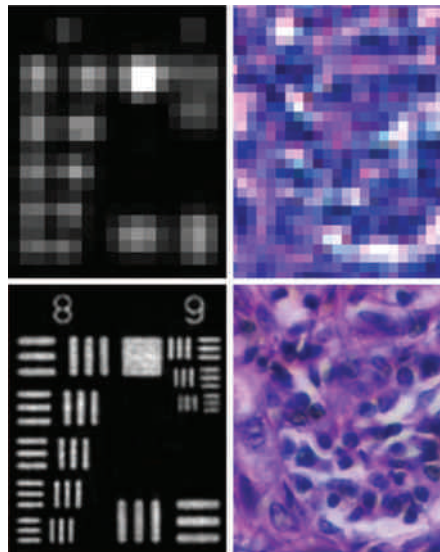
Raw data of a ×2 objective lens (top) before being rebuilt as a high-resolution image (bottom).

TinEye, which allow a user to upload an image to find out what it is and where you might find more like it. Balis demonstrates the system from his laptop; the image analysis takes about 10 seconds. "This is the simplest possible search structure for a two-dimensional image," he says. "It's a good quick-and-dirty tool."

The SVIQ software can help pathologists find sites of interest in the visually overwhelming landscape of a digital slide. Whereas C-Path is designed to make diagnostic decisions, SVIQ is a way to make digital pathology images more user friendly — and to extract information from them quickly. Jason Hipp, who works on pathology informatics at the National Cancer Institute in Bethesda, Maryland, uses SVIQ to quantify features over an entire slide. Instead of scoring the number of cells that are dividing in a few visual fields,

as pathologists do, he says, SVIQ can count every single dividing cell in the entire tissue slice. "When we have to count and measure, it's time consuming," he says. Using SVIQ could speed up the pathologist's work and provide more morphological data.

One of Hipp's research goals is to integrate SVIQ into the process of screening patients for cancer clinical trials. This often starts with a genetic test to find eligible patients. However, for tumours that are dispersed, it can be difficult to find enough cancerous tissue to perform the genetic screen, as the cancer's genetic signal can be lost in the noise from normal cells. It's also difficult to do genetic tests on samples taken by needle biopsies, which are less invasive but produce less tissue to work with, says Hipp. In these cases, pathologists have about 25 minutes after taking a biopsy from the freezer to identify and hand-dye cancerous portions of a tumour slice before it deteriorates. A laser is used to remove the undyed sections, leaving enriched cancer cells for genetic screening.

The SVIQ software can help to digitize this entire process. A pathologist finds an area of the image where there are cancer cells, uses SVIQ to highlight the rest of the cancer, and this map is then sent to the laser cutter. The process takes just 5 minutes.

At the Institute of Cancer Research in London, bioinformatics researcher Yinyin Yuan aims to map all the cell types in a tumour alongside their gene expression data. "The different cell populations in a tumour create a complex landscape that is an obstacle to accurate diagnosis," she says. A sequencing study that samples part of a tumour cannot capture the full picture: it blurs the role played by support cells and misses the heterogeneity of the cancer-cell population. These issues affect patients' prognoses and how they will respond to different kinds of therapy.

In 2012, Yuan developed software to classify the identity and distribution of each of the million or so cells from 300 whole-tumour slides of breast-cancer biopsies, and then integrated this with other '–omics' data. A human pathologist would take too long to go through so many cells in this manner, but a cluster of 100 computing cores, each with the power of a PC, can do the job overnight. The output from Yuan's software[3] is not an image, but data, analysed and collated with other data about the tumour. Yuan found that patients with immune cells that infiltrated the tumour had a better prognosis, and this prediction was strengthened when the image data were coupled with gene expression data. This result is not obvious to the eye of a pathologist staring at a slide and glancing at a list of gene expression data, but it becomes clear when Yuan's software analyses the image. Yuan is now expanding the project to study ovarian and lung cancer.

"What's happening in medical imaging is similar to what happened in astrophysics,"
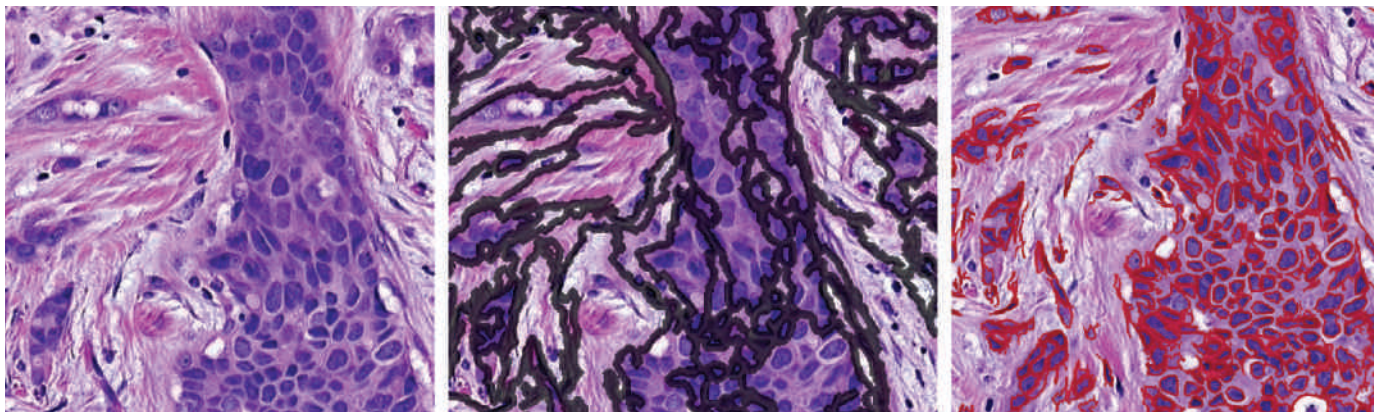
Breast cancer image at X200 magnification (left) is broken down into superpixels (black) by an algorithm before it predicts the patient's prognosis.

says Yuan. When astronomers got access to powerful telescopes and digital images, they didn't insist on counting every star — they let computers take over such tedious tasks. It should be the same with digital microscopes and cells, she says.

MATHEMATICAL LENSES

Using software to analyse digital images will make the pathologists' job easier. But it will also lead to new kinds of hardware. Today, creating digital microscopy images means scanning microscopy slides, and this is slow and expensive. Researchers are now building microscopes that can do both jobs, relying more on the power of software and less on lenses and other expensive hardware.

Slide scanners take multiple images, mechanically repositioning the slide under a microscope each time and then stitching the images together. "Mechanical scanning is slow," says Changhuei Yang, who develops microscope technologies at the California Institute of Technology in Pasadena. Because of their expense, these scanners are not typically found in community hospitals. Yang's solution is to increase the field of view and resolution of conventional microscopes. In July 2013, Yang hacked a low-resolution light microscope to create a high-resolution microscope with a wide field of view that can create whole-slide images with cheap hardware[4]. In conventional microscopes, low-power lenses provide a wide field of view at the expense of resolution; high resolution only occurs in small fields of view, hence the need for mechanical scanning. Yang's computational microscope can image an area as large as 120 square millimetres at a resolution of 0.8 micrometres; a comparable standard microscope offers a field of only 1.1 square millimetres at this resolution.

The trick is to use an array of light-emitting diodes programmed to sequentially illuminate the sample with three different colours of light from several different angles; the microscope records a picture each time. These images are then combined, picking apart the way the sample bent or changed the colour of the different light sources, to reconstruct a single, large-area, high-resolution image. Beck, who invented C-Path, says these images are comparable in quality to those made with expensive slide scanners. It's also more efficient, Yang says: "The lead time is shorter, so the number of samples a pathologist can examine can increase." Yang has just started up a company, Clearbridge BioPhotonics, based in Singapore, to commercialize the technology.

Some scientists are going one step further and building totally lens-free microscopes. Aydogan Ozcan, an electrical engineer at the University of California, Los Angeles, is developing microscopes that are basically just light-sensing electronic chips of the kind found in consumer electronics, but altered to cope with wet biological samples. The fancy part is Ozcan's software, which does the same thing as a physical lens: it transforms blurry interference patterns into focused images of cells. His compact microscopes[5] reveal the same details as those with lenses — those made with state-of-the-art chips have a resolution of hundreds of nanometres, clear enough to reveal the nuclei of cells. Sample preparation is similar to that for conventional microscopy.

The components needed to build these microscopes cost just a few dollars, and the calculations can be performed by the processors found in mobile phones. In fact, for demonstration purposes, Ozcan has built several microscopes attached to mobile phones. "We want to empower point-of-care offices or small clinics to work like a hospital lab," says Ozcan.

He is also experimenting with crowd-sourcing diagnostics. He uploaded the images of blood cells made by his microscopes to an online game (go.nature.com/mnmsmy) that teaches non-experts to recognize cells infected with malaria. The same images are shown to many different players, and by statistically combining the answers — after removing those clearly trying to upset the system — Ozcan's software generates the same diagnosis as professional pathologists 99% of the time[6]. The idea isn't to 'gamify' pathology — although the games might serve as training tools for medical students and lab technicians. Eventually, Ozcan says, smart software will be able to take over from human pathologists. In the meantime, says Ozcan, "I think we'll see hybrid modalities like this before machine learning takes over completely."

Innovations in computing are set to transform the field of pathology, says Alan Nelson, a physicist and chief executive of VisionGate, a company based in Phoenix, Arizona, that is developing three-dimensional imaging for the automated detection of cancer cells in sputum and blood. "A machine doesn't give an opinion — it can produce data and absolute diagnosis based on statistics," he says. The system could increase screening rates and help patients get the right treatment sooner.

Nelson previously was the lead inventor of the only automated cancer screening test currently on the market. His cervical cancer test, developed at his company NeoPath, received approval from the US Food and Drug Administration in 1996 and is now marketed by Becton Dickinson. This test uses a processor that is custom built for the specific problem of spotting cancer cells in pap smears. The machine is loaded with hundreds of slides that are scanned automatically all day.

Today's computers are capable of much more. Nelson says that microscopes aided by software are now showing biologists and doctors things they've never seen before. Ozcan's lens-free microscopes have revealed new patterns of helical motion in sperm, and Vision-Gate's three-dimensional images can show pathologists hundreds of previously unseen features. "We can see the texture of the inside of the nuclear surface of a lung cancer cell, and measure the length of the short arm of chromosome six," Nelson says. "My god, it's beautiful!" ∎

Katherine Bourzac *is a freelance science writer based in San Francisco, California.*

1. Beck, A. H. *et al. Sci. Transl. Med.* **3**, 108ra113 (2011).
2. Hipp, J. D. *et al. J. Pathol. Inform.* **2**, 13 (2011).
3. Yuan, Y. *et al. Sci. Transl. Med.* **4**, 157ra143 (2012).
4. Zheng, G., Horstmeyer, R. & Yang, C. *Nature Photon.* **7**, 739–745 (2013).
5. Greenbaum, A. *et al. Nature Meth.* **9**, 889–895 (2012).
6. Mavandadi, S. *et al. PLoS One* **7**, e37245 (2012).