



ILLUSTRATION BY THE PROJECT TWINS

A SHOUT-OUT FOR NOISE-REDUCTION TOOLS

Image de-noising approaches that can sharpen grainy or blurry microscopy pictures increasingly rely on machine learning. **By Amber Dance**

Saskia Lippens was satisfied with the images she was feeding into a program to identify borders between cells and organelles at the Flanders Institute for Biotechnology (VIB) in Belgium. So she hesitated when graduate student Joris Roels offered to use an algorithm to filter out the noise from the input micrographs, even though he said it would improve her border-finding results.

“An electron microscopist is always a little bit cautious when people talk about restoration of images,” says Lippens, director of the institute’s microscopy core facility in Ghent. “What if we’re changing data? What if we make mistakes?”

Roels, now a postdoctoral researcher at the VIB, convinced her by promising that a human biologist would always look over the computer’s shoulder, checking the images and

selecting the right de-noising approach. When Lippens tried to find those borders again using de-noised pictures, she was impressed with the improvement. “This was really a much better starting point,” she says.

Noise, put simply, is everything in an image that isn’t real signal. The weaker the illumination, the noisier the image, which explains the graininess of night-time mobile-phone selfies – not to mention low-light photomicrographs taken to protect fragile samples.

But no image is completely free of noise. “It’s always there,” says Michael Elad, a computer scientist at the Technion – Israel Institute of Technology in Haifa. To minimize it, researchers have long applied de-noising algorithms, the earliest incarnations of which were mathematical processes developed by computer scientists. “Then came the deep-learning era,”

says Elad. By passing the images to computers and allowing them to work out the best de-noising approach, researchers have begun to see striking results.

“It’s pretty magical,” says Loïc Royer, who works on image processing and light-sheet microscopy at the Chan Zuckerberg Biohub in San Francisco, California. But the magic does have risks: biologists must take care not to lose or muddle valuable signal.

From human to machine

Old-school, human-written algorithms still work well much of the time, and many have been built into popular image-processing environments, such as ImageJ, Fiji and MATLAB. Elad favours an approach called block-matching and 3D filtering, which groups together image sections that are similar in content and noise,

and filters noise from each group before reassembling the image. Carolina Wählby, director of the national SciLifeLab BioImage Informatics Facility at Uppsala University in Sweden, uses an algorithm called Top-Hat to clean up background noise in fluorescence micrographs and other images. (Top-Hat performs a mathematical transformation to remove overly bright or dark elements from an image.) “In many cases, something simple like that works really well,” she says.

Machine learning goes a step further, with the computer first learning how to de-noise one set of images, then applying what it has learnt to new data. “You skip the middleman, the mathematician,” says Royer.

There is one catch, however: the computer’s reasoning isn’t always apparent. “The learning algorithm is building a very complex black box that extracts the essence of what the images are about, and it just works,” says Royer.

There’s also the added computational cost, particularly during the training phase. Timely training might require computers linked to multiple graphics processing units or cloud-based servers. However, once the machine has finished this phase, researchers can usually de-noise images with a standard laptop, says Avinash Nehemiah, who manages computer-vision product marketing at MathWorks, the developer of MATLAB, in Natick, Massachusetts.

In ‘supervised’ learning-based approaches, the machine knows what it is looking for, because the user trains it with matching pairs of noisy and clean images. “Supervised, with the right data, will always give you the best results,” says Florian Jug, a computer scientist at the Max Planck Institute of Molecular Cell Biology and Genetics (MPI-CBG) in Dresden, Germany.

Supervised models get better as users provide more input pairs. But researchers don’t always have pre-cleaned images available. In that case, they can try algorithms that train themselves. According to Jun Xu, a computer scientist at Nankai University in Tianjin, China: “The new direction of this field is the development of self-supervised algorithms.”

Wave of the future

Some of the first entrants in this self-supervised category, Noise2Void and Noise2Self, were developed in the past few years. The developers assume that every pixel value is a bit off owing to random noise, but that nearby pixels should offer strong hints as to what the value should be, explains Royer, who co-developed Noise2Self. For each pixel, the machine uses the surrounding pixels to predict the proper value. It can then apply those parameters to new pictures.

“The noise doesn’t survive this process, because the noise is information that is only on that pixel,” says Royer.

Imaging specialists have already begun to improve on this technique. For example,

Noise2Self and Noise2Void fail if noise is not randomly distributed throughout the image. But Coleman Broaddus, a graduate student at MPI-CBG, tweaked Noise2Void to circumvent this issue. With his version, called StructN2V, users select a multi-pixel area that matches the size of the noisy bits of their image. Then, the machine-learning algorithm attempts to de-noise by predicting the centre value for that patch on the basis of surrounding pixels.

Imager beware

The output from such models, whether classical or based on machine learning, can look spectacular. But that doesn’t make it real. “There’s a trade-off,” says Nehemiah: the image is cleaner, but it also has been modified.

And the stronger the noise, the more likely it is that those changes will be significant. “When the noise becomes very strong – so strong that you hardly see the image – then the results are sort of hallucinations,” says Elad.

For example, in an attempt to eliminate blur from images scanned quickly, Wählby and a

“It’s really the biologist, the expert, who decides what you’re going to use and not use.”

student trained a machine with pictures of an empty electron microscopy grid that created a pattern of stripes. When they ran the trained model on new pictures, extra stripes appeared. “It learnt to find stripes, rather than to understand motion blur,” Wählby says.

The fix was to add more training data: specifically, a set of clean and noisy images without stripes. With that addition, the machine correctly learnt to remove the blur.

And some images are just too noisy to salvage. Rupali Mankar, who works with infrared imaging data at the University of Houston in Texas, says that she checks for this by taking multiple pictures of the same sample. If the output changes widely between images, she says, “it’s not a good signal, it’s just noise”.

Digital hallucinations and the like are relatively rare in Jug’s experience. When de-noising goes wrong, it’s usually obvious because images look blurry or weird, he says. “It is surprising how little these problems arise.” That said, Jug advises image analysts to include raw, noisy data in paper submissions or post the data on publicly available servers. That way, readers and reviewers can compare the data before and after for themselves. Many journals have teams that check images for mistakes in image manipulation, notes Kevin Eliceiri, a biomedical engineer at the University of Wisconsin–Madison.

Researchers should also take care to validate results using a secondary method, suggests

Wei Ouyang, a computer scientist at SciLifeLab at the KTH Royal Institute of Technology in Stockholm. For example, one might use a different imaging technique, such as a wider field of view, to confirm that the de-noised data make sense.

Picking the right approach

It was the biologist’s ‘human-in-the-loop’ evaluation that gave Lippens the confidence to de-noise her electron microscopy images. She, Roels and their collaborators designed an ImageJ plug-in called DenoisEM that’s equipped with eight classic de-noising approaches. Using a graphics processing unit for speed, microscopists can try different options and fiddle with parameters until they are satisfied with the results. “It’s really the biologist, the expert, who decides what you’re going to use and not use,” says Lippens.

For advice, researchers can try their local microscopy facility or an online community such as <http://forum.image.sc>. Mankar suggests that scientists who are new to imaging might also want to consider online classes, such as those offered by the online education firm Coursera, or a hands-on boot camp.

A growing collection of tools allows researchers to find and compare multiple de-noising approaches, and contribute new ones. For example, CSBDeep, developed by Jug and collaborators, is an online machine-learning toolbox that can be used with the Fiji image-processing environment or Python programming language. Likewise, Ouyang’s web app ImJoy offers a one-stop shop for test-driving multiple machine-learning methods.

Ouyang, Jug and others are also developing Bioimage Model Zoo. Microscopists will populate this repository with pre-trained machine-learning models for de-noising and other purposes, such as segmentation, which might allow users to skip the computationally costly training step. But borrowing another researcher’s model can be dangerous, says Jug, and it’s important that there is a close match between their data set and yours. A model trained to clarify microtubules, for instance, might fail when applied to pictures of nuclear membranes – or even to images of microtubules from a different microscope set-up. But if both signal and noise are similar, pre-trained models can work, says Lei Zhang, a researcher in low-level computer vision at the Hong Kong Polytechnic University.

In any event, today’s models and algorithms are not the last word in image de-noising, says Elad. “It’s an ever-running Olympics: everybody trying to beat everyone else.”

For researchers such as Lippens, the future – and photomicrography – has never looked clearer.

Amber Dance is a science journalist in Los Angeles, California.