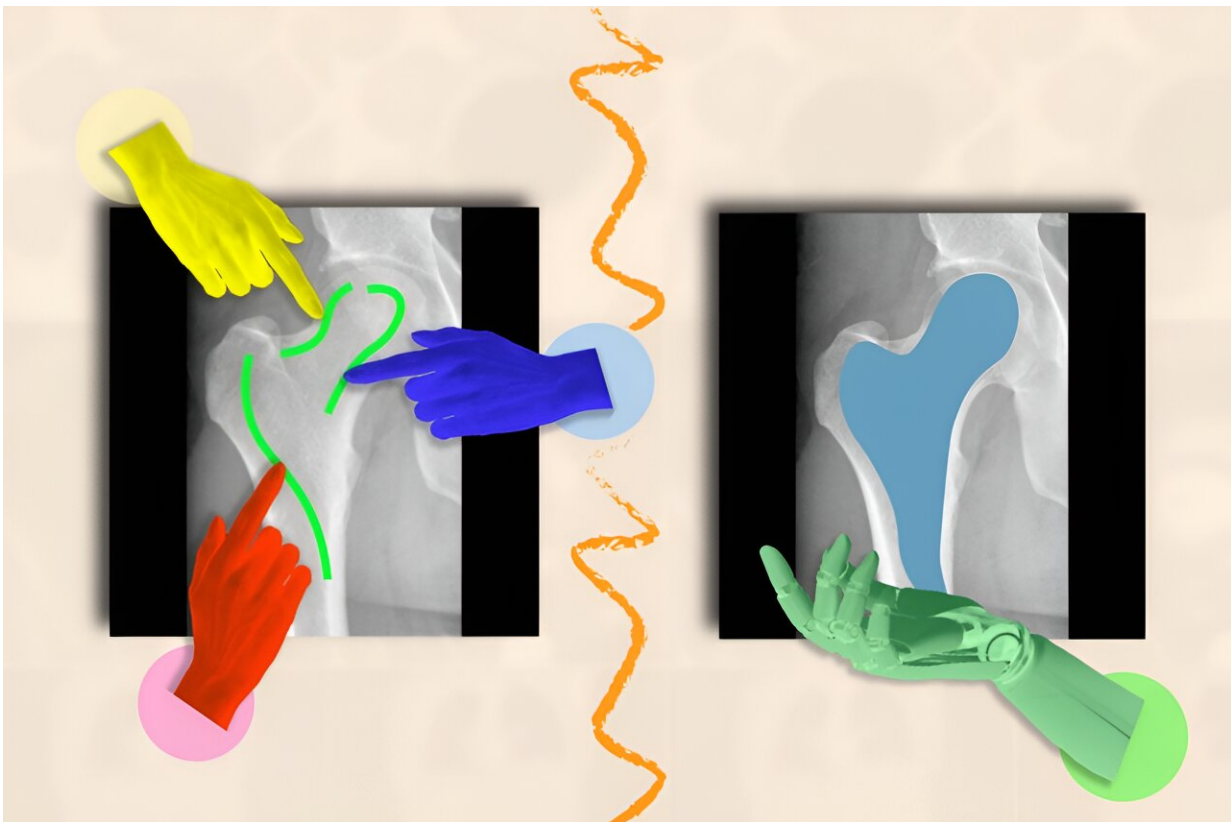# Interactive AI framework provides fast and flexible approach to help doctors annotate medical scans

September 9 2024, by Alex Shipps

ScribblePrompt's interface allows users to scribble across the rough area of a biomedical image they'd like segmented. They can also click on it or use a bounding box, and the tool will highlight the entire structure or background as requested. Credit: Alex Shipps/MIT CSAIL with elements from Pixabay, and the researchers.

To the untrained eye, a medical image like an MRI or X-ray appears to be a murky collection of black-and-white blobs. It can be a struggle to decipher where one structure (like a tumor) ends and another begins.

When trained to understand the boundaries of biological structures, AI systems can segment (or delineate) regions of interest that doctors and biomedical workers want to monitor for diseases and other abnormalities. Instead of losing precious time tracing anatomy by hand across many images, an artificial assistant could do that for them.

The catch? Researchers and clinicians must label countless images to train their AI system before it can accurately segment. For example, you'd need to annotate the [cerebral cortex](link) in numerous MRI scans to train a supervised model to understand how the cortex's shape can vary in different brains.

Sidestepping such tedious data collection, researchers from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts General Hospital (MGH), and Harvard Medical School have developed the interactive "ScribblePrompt" framework: a flexible tool that can help rapidly segment any medical image, even types it hasn't seen before.

Instead of having humans mark up each picture manually, the team simulated how users would annotate over 50,000 scans, including MRIs, ultrasounds, and photographs, across structures in the eyes, cells, brains, bones, skin, and more.

To label all those scans, the team used algorithms to simulate how humans would scribble and click on different regions in medical images. In addition to commonly labeled regions, the team also used superpixel algorithms, which find parts of the image with similar values, to identify potential new regions of interest to medical researchers and train

ScribblePrompt to segment them. This synthetic data prepared ScribblePrompt to handle real-world segmentation requests from users.

"AI has significant potential in analyzing images and other high-dimensional data to help humans do things more productively," says MIT Ph.D. student Hallee Wong SM '22, the lead author on a paper about ScribblePrompt and a CSAIL affiliate. The findings are published on the *arXiv* preprint server.

"We want to augment, not replace, the efforts of medical workers through an interactive system. ScribblePrompt is a simple model with the efficiency to help doctors focus on the more interesting parts of their analysis. It's faster and more accurate than comparable interactive segmentation methods, reducing annotation time by 28 percent compared to Meta's Segment Anything Model (SAM) framework, for example."

ScribblePrompt's interface is simple: Users can scribble across the rough area they'd like segmented, or click on it, and the tool will highlight the entire structure or background as requested. For example, you can click on individual veins within a retinal (eye) scan. ScribblePrompt can also mark up a structure given a bounding box.

Then, the tool can make corrections based on the user's feedback. If you wanted to highlight a kidney in an ultrasound, you could use a bounding box, and then scribble in additional parts of the structure if ScribblePrompt missed any edges. If you wanted to edit your segment, you could use a "negative scribble" to exclude certain regions.

These self-correcting, interactive capabilities made ScribblePrompt the preferred tool among neuroimaging researchers at MGH in a user study. 93.8 percent of these users favored the MIT approach over the SAM baseline in improving its segments in response to scribble corrections. As

for click-based edits, 87.5 percent of the medical researchers preferred ScribblePrompt.

ScribblePrompt was trained on simulated scribbles and clicks on 54,000 images across 65 datasets, featuring scans of the eyes, thorax, spine, cells, skin, abdominal muscles, neck, brain, bones, teeth, and lesions. The model familiarized itself with 16 types of medical images, including microscopies, CT scans, X-rays, MRIs, ultrasounds, and photographs.

"Many existing methods don't respond well when users scribble across images because it's hard to simulate such interactions in training. For ScribblePrompt, we were able to force our model to pay attention to different inputs using our synthetic segmentation tasks," says Wong. "We wanted to train what's essentially a foundation model on a lot of diverse data so it would generalize to new types of images and tasks."

After taking in so much data, the team evaluated ScribblePrompt across 12 new datasets. Although it hadn't seen these images before, it outperformed four existing methods by segmenting more efficiently and giving more accurate predictions about the exact regions users wanted highlighted.

"Segmentation is the most prevalent biomedical image analysis task, performed widely both in routine clinical practice and in research—which leads to it being both very diverse and a crucial, impactful step," says senior author Adrian Dalca SM '12, Ph.D. '16, CSAIL research scientist and assistant professor at MGH and Harvard Medical School. "ScribblePrompt was carefully designed to be practically useful to clinicians and researchers, and hence to substantially make this step much, much faster."

"The majority of segmentation algorithms that have been developed in image analysis and machine learning are at least to some extent based on

our ability to manually annotate images," says Harvard Medical School professor in radiology and MGH neuroscientist Bruce Fischl, who was not involved in the paper.

"The problem is dramatically worse in medical imaging in which our 'images' are typically 3D volumes, as human beings have no evolutionary or phenomenological reason to have any competency in annotating 3D images. ScribblePrompt enables manual annotation to be carried out much, much faster and more accurately, by training a network on precisely the types of interactions a human would typically have with an image while manually annotating. The result is an intuitive interface that allows annotators to naturally interact with imaging data with far greater productivity than was previously possible."

Wong and Dalca wrote the paper with two other CSAIL affiliates: John Guttag, the Dugald C. Jackson Professor of EECS at MIT and CSAIL principal investigator; and MIT Ph.D. student Marianne Rakic SM '22. Their work was supported, in part, by Quanta Computer Inc., the Eric and Wendy Schmidt Center at the Broad Institute, the Wistron Corp., and the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health, with hardware support from the Massachusetts Life Sciences Center.

Wong and her colleagues' work will be presented at the 2024 European Conference on Computer Vision and was presented as an oral talk at the DCAMI workshop at the Computer Vision and Pattern Recognition Conference earlier this year. They were awarded the Bench-to-Bedside Paper Award at the workshop for ScribblePrompt's potential clinical impact.

  More information: Hallee E. Wong et al, ScribblePrompt: Fast and Flexible Interactive Segmentation for Any Biomedical Image, *arXiv* (2023). DOI: 10.48550/arxiv.2312.07381

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology