Automated MRI image labelling processes 100,000 brain exams in under 30 minutes

July 22 2021

Credit: CC0 Public Domain

Researchers from the School of Biomedical Engineering & Imaging Sciences at King's College London have automated brain MRI image labeling, needed to teach machine learning image recognition models, by

deriving important labels from radiology reports and accurately assigning them to the corresponding MRI examinations. Now, more than 100,00 MRI examinations can be labeled in less than half an hour.

Published in *European Radiology*, this is the first study allowing researchers to label complex MRI image datasets at scale.

The researchers say it would take years to manually perform labeling of more than 100,000 MRI examinations.

Deep learning typically requires tens of thousands of labeled images to achieve the best possible performance in image recognition tasks. This represents a bottleneck to the development of deep learning systems for complex image datasets, particularly MRI which is fundamental to neurological abnormality detection.

Senior author, Dr. Tom Booth from the School of Biomedical Engineering & Imaging Sciences at King's College London said: "By overcoming this bottleneck, we have massively facilitated future deep learning image recognition tasks and this will almost certainly accelerate the arrival into the clinic of automated brain MRI readers. The potential for patient benefit through, ultimately, timely diagnosis, is enormous."

Dr. Booth said their validation was uniquely robust. Rather than evaluating their model performance on unseen radiology reports, they also evaluated their model performance on unseen images.

"While this might seem obvious, this has been challenging to do in medical imaging because it requires an enormous team of expert radiologists. Fortunately, our team is a perfect synthesis of clinicians and scientists," Dr. Booth said.

Lead Author, Dr. David Wood from the School of Biomedical

Engineering & Imaging Sciences said that "this study builds on recent breakthroughs in natural language processing, particularly the release of large transformer-based models such as BERT and BioBERT which have been trained on huge collections of unlabeled text such as all of English Wikipedia, and all PubMed Central abstracts and full-text articles; in the spirit of open-access science, we have also made our code and models available to other researchers to ensure that as many people benefit from this work as possible."

The authors say that while one barrier has now been overcome, further challenges will be, firstly, to perform the deep learning image recognition tasks which also have multiple technical challenges; and secondly, once this is achieved, to ensure the developed models can still perform accurately across different hospitals using different scanners.

Dr. Booth says that "this study was possible thanks to a very broad team of experts who are working on these challenges. There is a huge base of supporting organizers and facilitators who are equally important in delivering this research. Obtaining clean data from multiple hospitals across the UK is an important step to overcome the next challenges. We are running an NIHR portfolio adopted study across the UK to prospectively collect brain MRI data for this purpose."

More information: David A. Wood et al, Deep learning to automate the labelling of head MRI datasets for computer vision applications, *European Radiology* (2021). DOI: 10.1007/s00330-021-08132-0

Provided by King's College London