Journal: Neuroinformatics > Issue 1/2019

Authors: Ricardo Pizarro, Haz-Edine Assemlal, Dante De Nigris, Colm Elliott, Samson Antel, Douglas Arnold, Amir Shmuel

Abstract

Neuroimaging science has seen a recent explosion in dataset size driving the need to develop database management with efficient processing pipelines. Multi-center neuroimaging databases consistently receive magnetic resonance imaging (MRI) data with unlabeled or incorrectly labeled contrast. There is a need to automatically identify the contrast of MRI scans to save database-managing facilities valuable resources spent by trained technicians required for visual inspection. We developed a deep learning (DL) algorithm with convolution neural network architecture to automatically infer the contrast of MRI scans based on the image intensity of multiple slices. For comparison, we developed a random forest (RF) algorithm to automatically infer the contrast of MRI scans based on acquisition parameters. The DL algorithm was able to automatically identify the MRI contrast of an unseen dataset with <0.2% error rate. The RF algorithm was able to identify the MRI contrast of the same dataset with 1.74% error rate. Our analysis showed that reduced dataset sizes caused the DL algorithm to lose generalizability. Finally, we developed a confidence measure, which made it possible to detect, with 100% specificity, all MRI volumes that were misclassified by the DL algorithm. This confidence measure can be used to alert the user on the need to inspect the small fraction of MRI volumes that are prone to misclassification. Our study introduces a practical solution for automatically identifying the MRI contrast. Furthermore, it demonstrates the powerful combination of convolution neural networks and DL for analyzing large MRI datasets.