Exploring Extreme Edge Computing for Brain Tumor Classification in Magnetic Resonance Imaging with Tensor Decomposition

Junayd Lateef[1], Shahnawaz Syed[1], Niraj Anil Babar[2], Dr. Andreas Spanias[2], Dr. Julia Dietlmeier[3], Dr. Noel O'Connor[3]
[1] SCAI at Arizona State University (ASU), [2] SenSIP, School of ECEE at ASU, [3] Insight SFI Research Centre for Data Analytics at Dublin City University

IRES DCU Site: Sensor Information Processing & Machine Learning for Wearable Devices

Problem Statement

- Objective: Improve the overall performance of an AI Model for brain tumor image detection (accuracy, inference time, model parameters)
- Brain Tumor classification by specialists tend to take a long time due to the large number of factors that are needed to be analyzed
- Our previous model ANSA has good classification accuracy, however, inference speed and parameters used we not favorable
- Want to improve the speed while also keeping a high accuracy

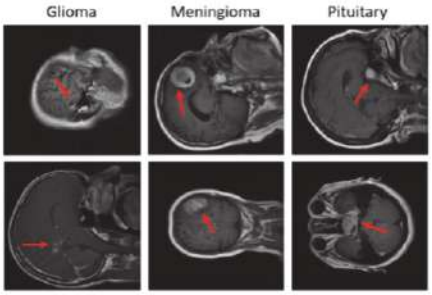

Fig 1. The three type of brain tumors being classified. [1]

Accuracy	Inf. time	Training time	Model parameter
93.312%	0.0025014	2049.7159 (ES)	5,471,875
93.637%	0.002921		5,559,443

Fig 2. Baseline model training runs of AlexNet (top row) and our ANSA model (bottom row)

Proposed Solution

- Utilized tensor decomposition to reduce the model parameters
  - Tucker Tensor Decomposition
  - Canonical Polyadic Tensor Decomposition
- Test and compare the performance of this model on multiple datasets to evaluate the generalizing capability of the model
- Further applications: Learn more about which parts of a model reduces its inference time and which reduces model parameters

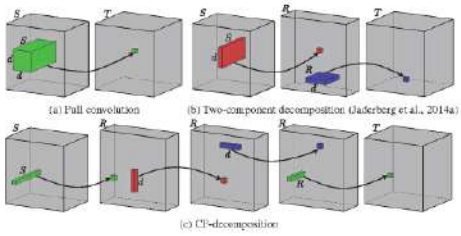

Fig 3. Different types of Tensor Decomposition. [4]

References
[1] J. Cheng, et al, "Retrieval of brain tumors by adaptive spatial pooling and fisher vector representation," PloSONE, vol. 11, no. 6, 2016.
[2] J. Cheng et al "Enhanced performance of brain tumor classification via tumor region augmentation and partition," PloSONE, vol. 10, no. 10, 2015.
[3] C. Billingsley, J. Dietlmeier, V. Narayanaswamy, A. Spanias, and N. E.O'Connor, "ANI2 normalized spatial attention network for accurate and fast classification of brain tumors in 2d t1-weighted contrast-enhanced images," 2023 IEEE ICIP, pp. 1895-1899, 2023
[4] G. Chenjin, J. Gildenblat, J. Kim, "Accelerating Deep Neural Networks with Tensor Decompositions", 2023.

http://sensip.asu.edu

IRES Site sponsored by NSF Award 2107439