

Graphics Processor Unit-Based Real-Time Reconstruction of Sensitivity-Encoded Radial Magnetic Resonance Imaging

Neha Bhatt

Asst. Professor, School of Computing, Graphic Era Hill University,
Dehradun, Uttarakhand India 248002

Abstract: Clinical uses of magnetic resonance imaging (MRI) have contributed to an increase in the difficulty of image reconstruction. However, a speedy computational method is required in order to diagnose and treat the condition. The advent of modern, competitive designs for graphics processing units (GPUs) has made high-performance parallel processing accessible and appealing to regular consumers. This has made it possible to compute massively parallel reconstruction problems at commodity costs. The reconstruction of an MRI using artificial intelligence (AI) places a larger load on graphics processing units (GPUs) due to the computations that must be performed in order to complete the task. This study was conducted with the intention of producing a comprehensive resource on the image reconstruction methods that may be used by the MRI research community. The resource will focus on GPU (Graphical Processing Unit) computation

Keywords: Graphics processing unit (GPU); magnetic resonance imaging (MRI); reconstruction

Introduction

When working in an MRI environment, medical staff would be subjected to high levels of magnetic field variations (up to 100 T/s in the range of frequencies 50-5000 Hz are generated) as well as powerful quasi-static magnetic fields. This could expose them to levels of magnetic field variation that could be harmful. Moving through magnetic fields may increase a person's exposure to extremely low frequency electromagnetic fields (ELF). This is because moving through magnetic fields can create rates of change in the magnetic field that can reach up to 20T/s in the frequency range of 0.1-50 Hz. Numerous studies [1-3] have been prompted by people's concerns about the biological implications of these interactions. When individuals in employment make continuous motions with their body, strong magnetic fields have the potential to cause huge induced currents [1]. A maximum permitted induced current density (J) of electric field in the body is often specified by both international and national rules governing exposure to ELF radiation. In order to accurately explain the boundaries of any discovered biological repercussions, it is essential to have the capability to monitor the induced electricity inside the body at a high spatial resolution. Because it is not possible to test these generated currents or electrical forces directly, the only method that can be used to determine exposure levels and ensure compliance with standards is computerised modelling

of the induced currents for specific applied fields. This is the only method that can be used. Examining these biological consequences via the use of numerical simulation is a strategy that is both time and money efficient. When attempting to estimate the size of the generated electric fields for simple or regular geometries or for conductivity distributions that are homogenous, an easy analytical procedure might be of great assistance. Even though there are many different numerical simulation packages available, such as Finite Difference Time Domain (FDTD) and Transmission Line Method (TLM), neither of these approaches is ideal for solving extremely low frequency problems, such as those associated with human motion and rotation near magnets [2]. Even though there are many different numerical simulation packages available, neither of these approaches is ideal for solving extremely low frequency problems. You might utilise a form of pseudo-static Finite-Difference (FD) approximation rather than FDTD. This approximation is based on the assessment of nodal potentials and connected admittances [3]. For the purpose of organising the data in preparation for calculation, this FD approximation makes use of a very large number of cubic cells. Because of the length of the calculation durations, putting such a computational instrument into practise might be challenging. It is important to calculate induced electric fields at various sites of movement at different periods in order to conduct an analysis of the movements of the body throughout the course of a certain amount of time. On a computer with two cores of processing power, a single motion calculation for a human head model with a precision of three millimetres and sixty-five hundred thousand voxels takes approximately five hours, and achieving convergence requires one thousand iterations. To investigate a typical body motion, such as the rotation of the head at an average angular velocity of $\frac{1}{2}$ rad/s for one second, however, around 45 simulation frames are required. Because of the amount of processing time needed to reproduce many of the normal body movements associated with occupational workers, doing research on the consequences of modifying MRI scanner hardware is not a practical option. The General Purpose Graphics Processing Units (GPGPU) hardware acceleration, on the other hand, is a solution that fits in well with the quasi-static FD approach. In this study, we demonstrate how a single graphics processing unit (GPU) from NVIDIA may be used to replicate the effect that a powerful magnetic field has on the behaviours of workers in a variety of different types of businesses. The first example takes around eight minutes to duplicate head rotation motions at four different magnetic field strengths, but the corresponding CPU version would take more than three hundred thirty-six minutes to complete the task. The end result is a decrease in processing time that is more than 40 times faster. In the second illustration, a person is seen being dragged towards a 7T magnet. In this particular project, the torso model consists of close to 8 million individual cells and must be completed using 10,000 iterations. It is possible to get a speedup of increase to 27 times in comparison to the CPU.

Rapid or Sparse MRI

Researchers have been looking for feasible alternatives to increase the speed of MRI by sampling the k-space at a rate lower than the Nyquist rate, while maintaining the same level of picture quality despite the inherent limitations of the technique. However, in order to violate the Nyquist requirement for signal/image sampling, undersampling in k-space that is caused by periodic interleaving is necessary. Aliasing artefacts will be produced in the reconstructed signal or picture as a consequence of this. According to the Shannon-Nyquist theorem, which is generally followed by all signal or picture collecting systems, the sampling frequency must be at least twice as high as the highest frequency contained in the signal (the Nyquist rate). This is the minimum requirement for the sample frequency to meet. This approach, which converts analogue signals to digital ones and is used in every piece of consumer gear, every piece of medical equipment, and every radio receiver on the market

today, results in an enormous amount of data being dumped. As long as the natural signals or pictures are represented on an appropriate basis, compression may be applied to nearly any sort of natural signal or image. This indicates that just a few number of non-zero big magnitude coefficients need to be employed in order to accurately characterise the energy of the whole signal or picture. After applying the appropriate transformation to the signal or picture, the vast majority of the coefficients may be ignored. Lossy compression methods, such as JPEG-2000, are based on this basic idea as their foundation. In other circumstances, the difficulty is in simply obtaining access to the relevant information. The above raises an obvious follow-up question: is it not possible for us to reconstruct a signal or picture from just a few encoded measurements if we gather all of the samples and then toss away the majority of them during reconstruction??

- i) It is suggested to correctly recover a single-slice MR image from significantly undersampled observations using a Hidden Markov Model-based wavelet support detector..
- ii) For the purpose of estimating multi-slice MRI missing data in k-space, a unique 3D interpolation approach is presented..
- iii) To take advantage of the information redundancy in multi-slice MRI data at the wavelet levels, we designed a forest sparsity based joint CS reconstruction model.
- iv) We developed a completely calibration-free joint CS reconstruction model that can take use of the wavelet and spatial sparsity of multi-slice pMRI data, and we validated and tested it on real-world multi-slice pMRI datasets. For the purpose of establishing the suggested CS reconstruction model's clinical viability, it is also implemented in a CPU-GPU setting.

Literature Review

Chong Ma ET.AL.,(2022) Because of the high risk associated with brain tumours, receiving a timely diagnosis and treatment is of the highest significance. When it comes to segmenting the cancer sub-regions of the multi-modal brain tumour MRI, the performance of the 3D convolutional neural network is superior to that of the conventional 2D convolutional neural network. The method, taken as a whole, makes it easier to extract more contextual information, which, in turn, helps with the accurate segmentation of the brain tumour into its component pieces. The author of this study suggests employing group convolution rather than traditional convolution as a means of significantly cutting down on the amount of memory that is required in order to optimise the flow of information between the many components that are connected by the multi-fiber unit. The experiment uses the cross-GPU synchronisation normalisation strategy in order to remedy the subpar segmentation effect generated by the 3D convolutional neural network as a consequence of the low batch value and the inability to forecast the gradient direction. This is done in order to remedy the subpar segmentation impact produced by the 3D convolutional neural network. This is done in order to fully grasp the advantages that may be gained from the computer hardware that is now accessible. In the end, we changed the activation function that had been used before to one that was superior in terms of its ability to segment. The experiment described in this paper is verified by utilising the publicly available dataset of brain tumours known as Brats2018. The average Dice value of the revised model is 89.79% for the overall brain tumour region; this value is 85.59% for the area that contains the tumour core; and this value is 79.83% for the region that contains the cerebral tumour augmentation. When compared to earlier iterations of segmentation models, the training period for this experiment, which lasted 321.89 minutes, is relatively short.

Chi Zhang et.al.,(2020) Reconstruction accomplished via the use of physics-guided deep learning (PG-DL) has developed into a powerful method for rapid MRI. However, widespread use of PG-DL on 3D non-Cartesian MRI is still in its infant stages. This is primarily because of limitations in GPU hardware. Within the scope of this investigation, we use PG-DL to a large-scale 3D kooshball coronary MRI and make use of a wide range of memory-efficient techniques. To begin, we implement a recently proposed strategy for keeping GPUs on a single unrolled stride at all times. After that, a Toeplitz approach is used in order to properly depict the multi-coil encoding operator. Then, we get rid of the requirement for coil compression by distributing the data consistency operations that require the most memory across several GPUs, which enables iterations of the conjugate gradient. In conclusion, mixed-precision training is used in order to further reduce the need of using one's memory. Our research has shown that this combination of techniques produces the best results when it comes to training high-quality PG-DL reconstruction for 3D kooshball trajectories. This discovery suggests that this combination of methodologies is the best option for teaching high-quality reconstruction.

Dushyant Sahoo et.al (2018) In recent years, there has been a rise in curiosity about resting-state functional magnetic resonance imaging (MRI), a technique that enables researchers to evaluate the functional connectivity throughout the whole of the brain. In order to examine the dynamics of functional connectivity, this approach is being used more and more, particularly for the goal of identifying individual biomarkers. On the other hand, there are major barriers to overcome as a result of the high amounts of background noise that are present in each every fMRI scan. In this article, we propose using Granger causality patterns in order to examine the dynamics of fMRI among people. This approach is able to simultaneously derive population causality patterns that are more resistant to noise while also capturing the organisation of the brain at the individual level. We construct an efficient method for finding shared causation patterns by evaluating the rs-fMRI data of one hundred people who are not linked to each other who took part in the Human Connectome Project. We then demonstrate the method's effectiveness..

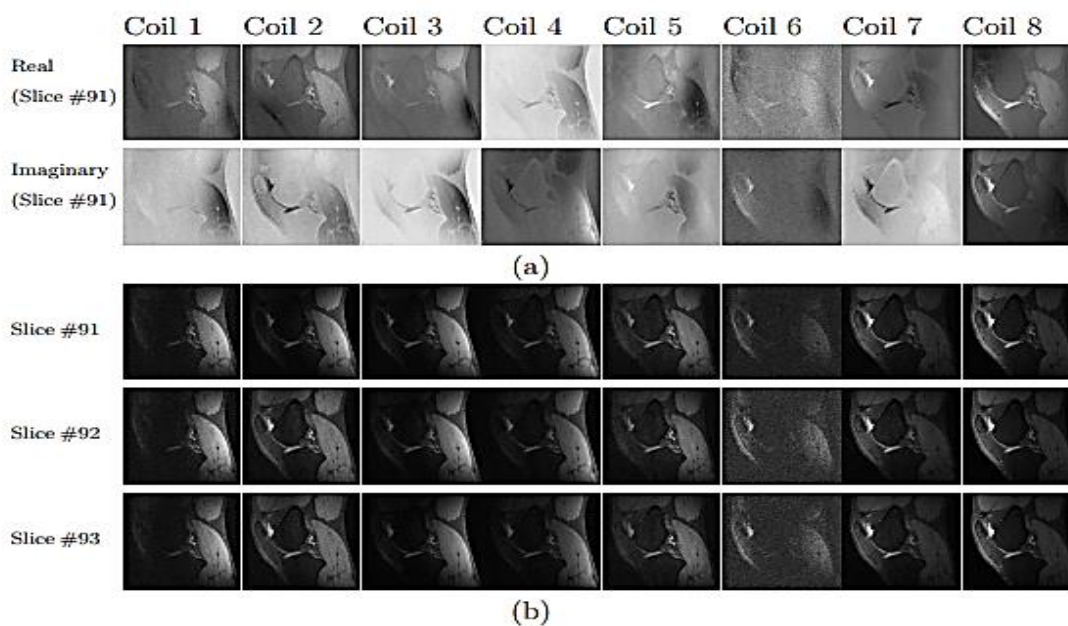


Figure 1: Multi-slice pMRI data. (a) Real and imaginary parts of complex Knee MRI data. (b) Magnitude of complex multi-slice Knee pMRI data (Source: www.mridata.org)

Proposed Method

By capitalising on similarities that occur simultaneously in slice and coil direction of multislice pMRI data, the fundamental objective of our research is to improve the quality of parallel MRI reconstruction from severely undersampled observations. This will be accomplished by using multislice pMRI data. The method that has been developed involves two stages: (a) k-space interpolation, and (b) calibrationless joint sparsity. These stages are used to reconstruct CS pMRI without the need of calibration. When it comes to interpolation, we make advantage of k-space redundant data along the slice-select axis. In addition, there are overlaps inside only one coil, across coils, and between slices of the 2D data. These overlaps may be found in all three locations. It is feasible to analyse tree sparsity by transforming coil data into an image and representing it using the wavelet domain. This may be done in one of two ways. A forest may be represented as a collection of wavelet trees that are connected to one another and consist of different coils and slices.

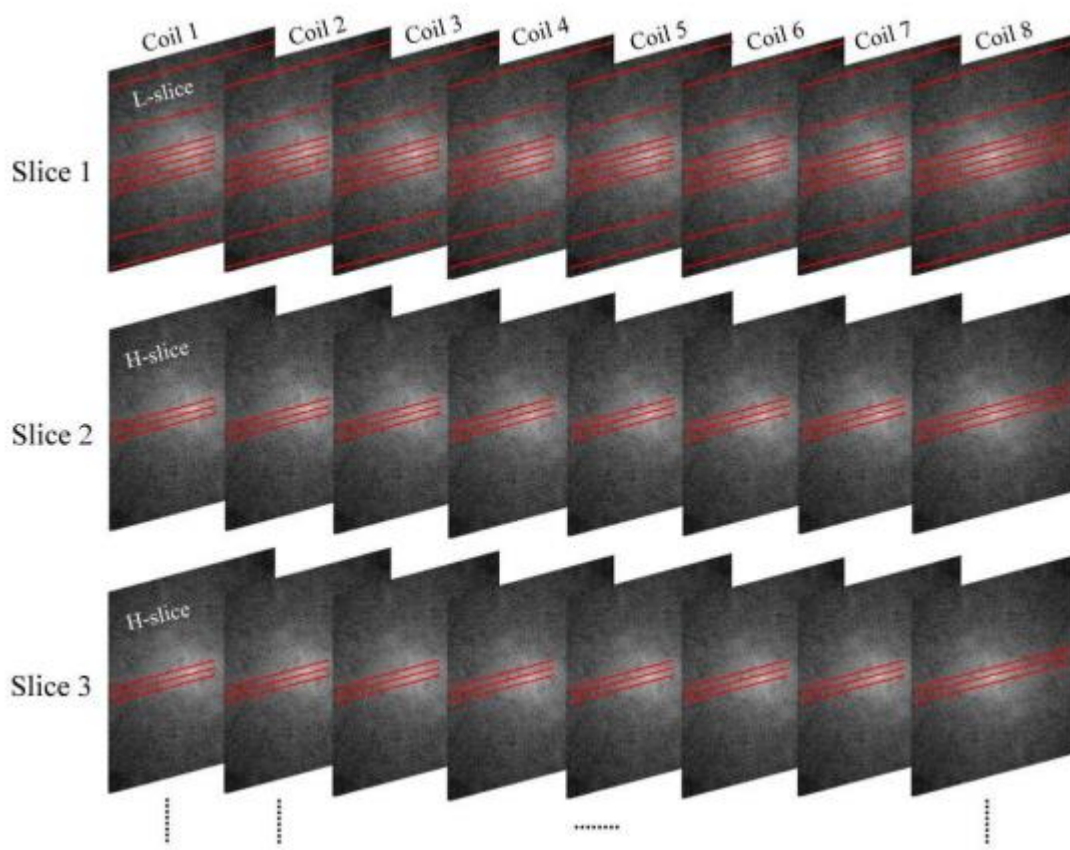


Figure 2: Graphical representation of the propose non-uniform undersampling scheme for multi-slice pMRI data

These data sets are comprised of MR images that are of a complex character. In Dataset I, there are eight channels of genuine knee data that were sampled on a 320320172 Cartesian grid. In Dataset II, there are fifteen channels of real knee data that were sampled on a 76877031 grid. Both of the datasets, which contain authentic k-space information, are subjected straight away to undersampling. We handle the real and imaginary portions of the data individually during the iterative reconstruction. This results in a large collection of reconstructed pictures that can be used to assess the correctness of the approach and compare

it to other methods. Real MR images are complicated data, thus we process them separately during the iterative reconstruction. At each of the four levels of decomposition, we put into action a sparsifying transformation that makes use of the Daubechies wavelet (db2). Because the majority of MR images are piece-wise smooth, the 'db2' wavelet could be a good choice. Higher order wavelets require more computation, but they do not significantly improve the quality of the reconstructed image. On the basis of study and testing [18, 52], the parameters for regularisation have been fixed in stone. These parameters include $\lambda_1 = 0.001$ and $\lambda_2 = 0.035$. The SNR, MSSIM, and FSIM of Dataset I reconstructions are shown in Figures 5.6 for a range of undersampling percentages. According to our findings, the performance of ESPIRiT is superior than that of SPIRiT and CS-SENSE, which are two additional auto-calibration approaches. The recommended method offers an advantage over the ESPIRiT in terms of performance since it is superior by a margin of 1.5 dB on average, the proposed technique could also end in an increase of 0.5–1 dB in SNR. Alterations of a positive nature may also be seen in the FSIM. Despite this, both the PLORAKS and the proposed offer superior performance in terms of how similar their structures are to one another.

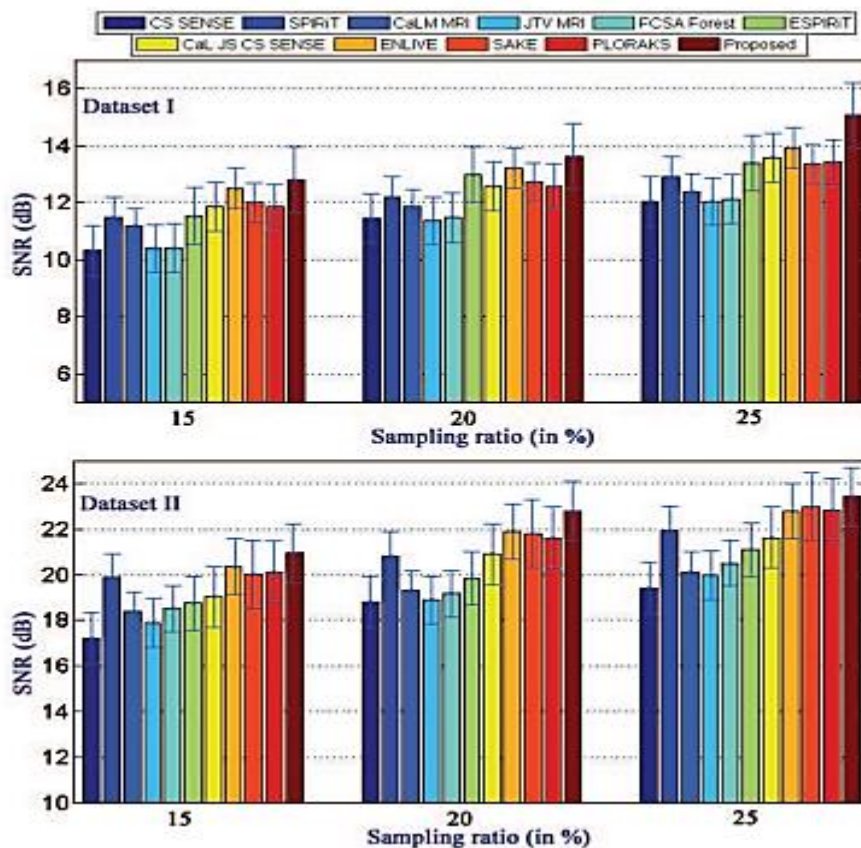


Figure 3: Reconstruction performance in terms of SNR for different datasets under different undersampling ratios

Parallel Implementation in CPU-GPU Environment

To solve the overall reconstruction problem, we break it up into smaller, more manageable pieces that can be run in parallel on multi-core machines. For example, in order to perform the 58 CS reconstruction problems in parallel, Dataset I (320 320 8 172) was split into two halves. First and last parallel subproblems take into account data of size (320 320 8 2), while the remaining 56 parallel CS reconstruction subproblems take into account data of size (320 320 8 3). These 4D subproblems are simultaneously tackled by multi-core machines. Using

distributed shared memory, each concurrent task performs computationally expensive operations locally, such as forward/backward Fourier and wavelets modifications, matrix-matrix/vector multiplications, and so on.

Table 1: Comparison of the computational cost in terms of the amount of time spent by the central processing unit (in minutes) for sequentially and parallel versions of the proposed approach with an under sampling ratio of 20%

Dataset	Sequential implementation	Parallel implementation
Dataset I	134.4	2.51
Dataset II	178.2	3.17

Graphics processing units with a high number of cores. This approach considerably cuts down the amount of time necessary to calculate the answer by virtue of the fact that many operations included inside each subproblem may make use of GP-GPU hardware in parallel. Programming on multi-core central processing units (CPUs) as well as graphics processing units (GPUs) have both made use of the MATLAB Parallel Computing Toolbox. The results of the computational analysis are shown in Table 1 for both datasets. We have found that the reconstruction process may be finished in a matter of minutes when a parallel incorporate-core and GP-GPU are used in conjunction with the strategy that has been presented. We only evaluated the recommended way to evaluate the practical practicality of GPU creation in terms of processing time. This is because the suggested approach is both one of the fastest algorithms in serial construction and the most exact of all of those that were researched for comparison. Consequently, it was the only one that we investigated. We come to the conclusion that the proposed method, if it has the appropriate parallel computing features, has the potential to provide clinically-viable reconstruction outcomes within a period that is clinically sensible..

Conclusion

Within the scope of this investigation, we propose a novel approach to the calibration-free CS-based reconstruction of pMRI. Because we have tested the proposed method using actual MRI datasets from the real world and compared it to the most advanced computer science-based pMRI algorithms, we are confident that it is successful. Experiments have proven that it is superior to other approaches in terms of both the quantity and quality of the results it produces. We have developed the approach using parallel computing on multi-core central processing units (CPUs) and general purpose graphics processing units (GP-GPUs) in order to further study the computational time and clinical feasibility of the method. In as little as two to three minutes, it's possible to successfully recreate clinical information by employing graphics processing units (GPUs). It has been shown that the calibration-less CS pMRI reconstruction approach that was developed is not only effective but also feasible.

Reference:

1. C. Zhang *et al.*, "Distributed Memory-Efficient Physics-Guided Deep Learning Reconstruction for Large-Scale 3d Non-Cartesian MRI," *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, Kolkata, India, 2022, pp. 1-5, doi: 10.1109/ISBI52829.2022.9761485.
2. R. Islam, M. S. Islam and M. S. Islam, "GPU based Real Time Reconstruction of Compressed Sampling MRI," *2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)*, Langkawi, Malaysia, 2020, pp. 35-39, doi: 10.1109/CSPA48992.2020.9068722.
3. D. Sahoo, N. Honnorat and C. Davatzikos, "GPU accelerated extraction of sparse Granger causality patterns," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 2018, pp. 604-607, doi: 10.1109/ISBI.2018.8363648.
4. S. B. Dinh, N. H. Le and H. M. Nguyen, "Accelerating GenSeT Reconstruction for Sparsely Sampled DCE-MRI with GPU," *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Louisville, KY, USA, 2018, pp. 29-34, doi: 10.1109/ISSPIT.2018.8642674.
5. T. Nguyen-Duc and W. -K. Jeong, "Compressed sensing dynamic MRI reconstruction using multi-scale 3D convolutional sparse coding with elastic net regularization," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 2018, pp. 332-335, doi: 10.1109/ISBI.2018.8363586.
6. R. S. Dehal, C. Munjal, A. A. Ansari and A. S. Kushwaha, "GPU Computing Revolution: CUDA," *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, Greater Noida, India, 2018, pp. 197-201, doi: 10.1109/ICACCCN.2018.8748495.
7. S. Saxena and S. Shama, "Brain Tumor Segmentation by Texture Feature Extraction with the Parallel Implementation of Fuzzy C-Means using CUDA on GPU," *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Solan, India, 2018, pp. 580-585, doi: 10.1109/PDGC.2018.8745726.
8. O. Zavala-Romero, A. Meyer-Baese and M. B. I. Lobbes, "Breast lesion segmentation software for DCE-MRI: An open source GPGPU based optimization," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 2018, pp. 211-215, doi: 10.1109/ISBI.2018.8363557.
9. Q. Meng and Z. Jin, "Online dynamic magnetic resonance imaging based on pseudo-polar sampling and GPU acceleration," *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Shanghai, China, 2017, pp. 1-5, doi: 10.1109/CISP-BMEI.2017.8302180.
10. T. -W. Huang *et al.*, "Age estimation from brain MRI images using deep learning," *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Melbourne, VIC, Australia, 2017, pp. 849-852, doi: 10.1109/ISBI.2017.7950650.
11. H. -Y. Tsai, Z. Hanyu, C. -L. Hung and H. -M. Chen, "GPU-based Gray-Level Co-occurrence Matrix for Extracting Features from Magnetic Resonance Images," *2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC)*, Exeter, UK, 2017, pp. 391-396, doi: 10.1109/ISPAN-FCST-ISCC.2017.80.

12. H. -Y. Tsai, H. Zhang, C. -L. Hung and G. Min, "GPU-Accelerated Features Extraction From Magnetic Resonance Images," in *IEEE Access*, vol. 5, pp. 22634-22646, 2017, doi: 10.1109/ACCESS.2017.2756624.
13. X. Wang and Z. Jin, "Online dynamic magnetic resonance imaging based on radial sampling and GPU acceleration," *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Datong, China, 2016, pp. 581-585, doi: 10.1109/CISP-BMEI.2016.7852777.
14. T. M. Quan and W. -K. Jeong, "Compressed sensing reconstruction of dynamic contrast enhanced MRI using GPU-accelerated convolutional sparse coding," *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, Czech Republic, 2016, pp. 518-521, doi: 10.1109/ISBI.2016.7493321.
15. M. Sabbagh, M. Uecker, A. J. Powell, M. Leeser and M. H. Moghari, "Cardiac MRI compressed sensing image reconstruction with a graphics processing unit," *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, Worcester, MA, USA, 2016, pp. 1-5, doi: 10.1109/ISMICT.2016.7498891.