



脳MRIデータ前処理の高速化 ～大規模計算による脳機能解明の実現に向けて～

研究・開発機関 : [GMO学術サポート&テクノロジー株式会社](#)
 利用施設 : スーパーコンピュータ「富岳」
 計算規模 : 1410枚の脳MRI画像データ前処理 17時間20分
 利用ソフトウェア : FreeSurfer (構造MRI解析ソフトウェア)

Before

- 脳画像の前処理は計算が複雑で時間がかかり、一般的なPCでは1枚の脳画像処理に半日から1日要していました。
- そのため、大量の画像処理には現実的でない長時間の計算が必要になり、大規模な脳画像データの高速処理が必要な神経科学研究などへの活用は困難でした。

After

- スーパーコンピュータ「富岳」を使用し、パラメータ最適化を実施することにより、脳画像の前処理の計算時間を大幅に短縮しました。
- 1410枚の脳画像の前処理を並列処理により約17時間で完了可能となり、神経科学研究などへの応用に大きな進展をもたらすことが可能になりました。

背景と目的

近年、神経科学分野では複雑な脳活動を解析するために、大量の脳画像データを集積し、大規模データ解析を行うことが必要となっています。特に、脳の機能や構造を理解するために行われる磁気共鳴画像(MRI)の解析は、非常にデータ量が多く、複雑な計算を必要とします。

脳画像解析を行うためには、被験者の個人差の影響を排除するための標準化の画像処理が必要です。これは頭蓋骨のような脳以外の組織を画像中から除去することや、画像のむら補正、灰白質(中枢神経系の中で神経細胞が密集する部分)と白質(神経細胞がなく、有髄神経線維ばかりの部分)の区別といった処理です。これにより、個々人の多様な脳形態を集団解析可能な標準的な脳画像に変換します。

従来の方法では、一般的なコンピュータを使用して1枚の脳画像の前処理に半日から1日かかることもあり、大量の画像データに対して適用する場合、現実的ではないほど時間がかかるという問題がありました。

GMO学術サポート&テクノロジー株式会社はこの課題を解決するため、スーパーコンピュータ「富岳」を利用した脳画像データの高速前処理方法を開発することを目的としました。「富岳」の大規模

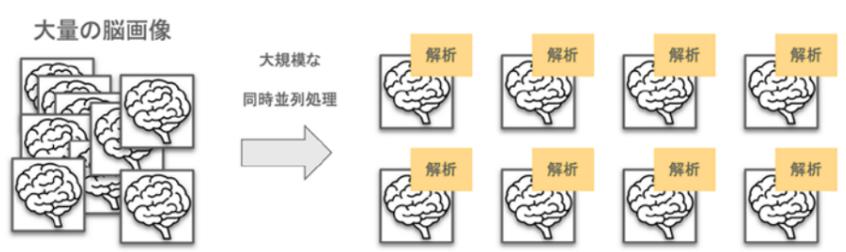


図1 大量の脳画像の並列処理

並列計算能力を活用することで、従来の方法よりもはるかに短い時間で脳画像の前処理を行うことができると期待でき、大規模なデータ分析が可能になります(図1)。

この研究は大規模な脳画像データの分析を通じて、特に神経科学分野において、脳の機能や病気の理解を深めるための基盤となると期待されます。また、スーパーコンピュータを利用したデータ処理の手法は、他の科学分野においても応用可能であり、研究の幅を広げる可能性を持っています。

利用成果

利用したデータとソフトウェア

本研究では脳MRIデータとして、SRPBS Multidisorder MRI Dataset^[1]で提供されている、身体の解剖学的な構造が見やすい等の特徴があるT1強調像を用いました。並列化と処理時間の関係性を正確に評価するため、1被験者のデータを1410個複製し1410人分に当たるデータセットを作成し、評価に用いました。また、脳画像の前処理のためのソフトウェアは脳画像解析で広く用いられるFreeSurfer^[2]を使用し実施しました。

解析手法

1. ノード*上でのFreeSurfer処理の最適化

FreeSurferを用いた脳画像の処理速度を向上させるために、シングルノード上でのスレッド並列化の度合いを変化させて評価し、並列数の最適化を行いました。結果として「富岳」のノードあたりのメモリ容量(32GB)に対し8被験者分のデータを並列解析することが最適であることが明らかになりました。

ノード*:CPUとメモリからなる管理単位、「富岳」ではノード当たり48CPUコア

2. マルチノード処理への拡張

シングルノードでの処理を最適化した後、処理をマルチノードへ拡張し、大量の脳画像データを並列で処理しました。1ノードで8被験者分のデータを解析するために、1410個の脳MRI画像を8人分ずつにまとめてバルクとし、全体で177バルクのデータを並列処理しました(図2)。

3. 実行時間と並列化率の評価

全体の実行時間と並列化率を評価し、効率的な並列処理が可能であることを確認しました。これにより、1410個の脳画像データの前処理をわずか17時間20分で完了することができました。

この研究は、大規模な脳画像データの高速処理と神経科学研究への応用に大きな影響を与えられます。このとき並列化効率(アムダールの法則)は99.98%(理論上の最大値は100%)でした。

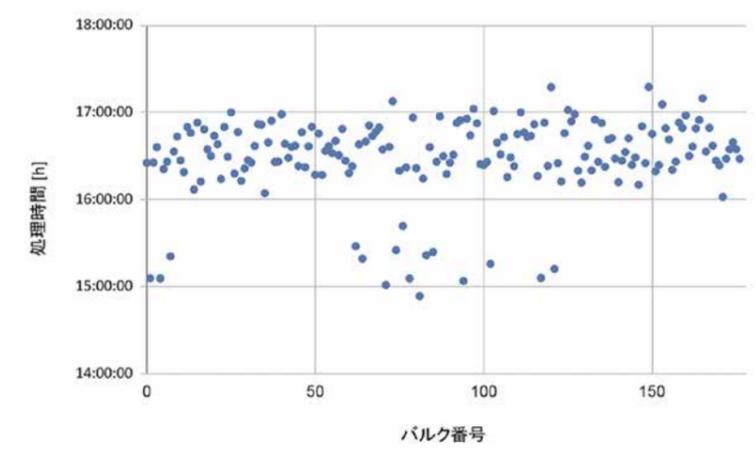


図2 バルク番号ごとの処理時間

結論

「富岳」の177ノードを用いて並列計算を実行することによって、約17時間という短時間で脳画像前処理を実現することができました。

この結果は、これまで計算時間の問題で不可能だった大規模データの分析を可能にします。また、脳画像前処理の精度を向上させるための事前パラメータ調整について、処理を並列実行することで最適パラメータ探索を高速化することも可能であると考えられます。

出典: [1] Tanaka SC, Yamashita A, Yahata N, Itahashi T, Lisi G, Yamada T, et al. A multi-site, multi-disorder resting-state magnetic resonance image database. Sci Data. 2021 Aug 30;8(1):227.
 [2] Fischl B. FreeSurfer. Neuroimage. 2012 Aug 15;62(2):774-81.